U. S. AIR FORCE
PROJECT RAND

THE THEORY OF LINEAR INEQUALITIES

Translated by
D.R. Fulkerson

T-22

7 March 1952

Assigned to

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

Best
Available
Copy

INTRODUCTION

§ 1. Survey

1. Origin. In the sign rule of Descartes and its generalizations, which has been the subject of investigation by many others, particularly Laguerre, the number of changes of sign plays an important role. Fekete and, after him, Pólya, had the important idea of investigating in general how the number of sign changes varies when the finitely many numbers of the sequence are considered as variables and subjected to a linear transformation.

The central problem turned out to be that of determining all those linear transformations which never increase the number of sign changes. The matrices of such transformations will be termed variation decreasing.

In 1930 Schönberg solved this problem for the case that the rank of the matrix of the transformation is equal to the number of independent variables. For this purpose he introduced the notion of a minor-definite matrix. With the help of a natural sharpening of this notion I solve the problem for the general case. In this case, as in the one Schönberg treated, it is sufficient for the decision to know the signs of all the minors of the coefficient matrix.

2. Central questions. The deeper basis of this fact is that the knowledge of these signs also suffices for the answering of many more general questions: namely for the deduction of

explicit conditions for the solvability or unsolvability of a system of linear inequalities. In §§3-12 we develop a general theory of such systems and their solvability and then deal independently with the variation decreasing matrices in §§13-14.

To the best of my knowledge, Fourier first treated linear inequalities systematically. Later, independently of him, Gordan and Minkowski successively turned to this subject, and in this century about 30 papers, listed in the bibliography, have been published on linear inequalities. A shorter and somewhat incomplete summary of the historical development can be found in the work of Ruth Stokes. Dines and McCoy give an extensive bibliography and historical notes in their combined work.

3. <u>A sketch of the investigation.</u> Statements about equation systems can be divided into the quantitative theory of the explicit representation of solutions by determinants (e.g. Cramer), and the qualitative theory of dependence, rank, etc. In the same way we divide our investigation into two parts (§§5-10 and §§11-14).

We first consider homogeneous systems and for this purpose define the basis (§5), prove its existence (§6), state it explicitly (§7), and consider it in more detail (§8). Then we obtain a parametric representation of all solutions which can be further reduced for the "fixed" system (§9), whereupon we are also in a position to solve the inhomogeneous system (§10).

Then we proceed to the equivalence and solvability problems. By determining the rank of the "minimal system" (§11) we obtain a criterion for unsolvability (§12), from which two theorems and

their corollaries follow: the transposition theorem (§13) and simplex theorem (§14).

In general we have presented the analytic theorems together with the corresponding geometric statements. The questions we pose deal with the group belonging to affine geometry.

4. Results. Because of the many interrelations between statements in the subject dealt with, it will often be difficult to decide when a theorem is really new.

Except for the presentation itself, which I hope appears in a new correction, I would like to emphasize the following places where progress is made.

The idea of considering inequality systems for all possible sign ranges (sign combinations) is perhaps used in the present work for the first time. The discovery of a basis for this most general case saves many special considerations.

Besides, the so-called combination principle can be understood and formulated only in this way. From this principle there arise proofs of the transposition and simplex theorems, and the transposition theorem obtains a particular general form.

5. Notation and terminology. The numbers represent the paragraphs in which the notion first appears.

alternate $\pm$	70, 93	closed	21
basis B	14, 17, 33, 45	coefficient range E_i, W	31
central part G	45	column combination Ai	16
central projection	65	column number m	16

6. Theorems and formulas. The numbering is as in 5.

Theorems:

A1	17	C1	34	E1	30	G1	74
A2	18	C2	38	E2	50	G2	74
A3	88	C3	39	E3	59	G3	75
A4	88	C4	44	E4	65	G4	75,78
A5	89	C5	58			G5	76,77
A6	90						
B1	70	D1	68	F1	22		
B2	91	D2	69	F2	22		
B3	92	D3	71	F3	24		
B4	92	D4	71	F4	80		
B5	93	D5	72	F5	81		
B6	95	D6	73,87	F6	83		
B7	96	D7	80	F7	85		
				F8	85		

Formulas:

§2. History and Bibliography[1]

7. Fourier. The beginning of interest in linear inequalities as a special kind of research goes back to Fourier, whose penetrating view first saw the dormant possibilities and who first thought about a systematic theory of linear inequalities, indeed **of general inequalities.**

It had been noticed before him that inequality conditions underlie the principles of mechanics of constrained systems, and Gauss had sought to avoid them by his principle of least constraints. Now in the case of statics these conditions assume simpler and often linear forms. Fourier was led from this to our problem, as he had always sought to attack concrete relations by abstract means.

Fourier, however, did not go deeply into the problem. His solution method is the successive one (later called reduction[2]), along with an interesting geometric interpretation which reveals, for the first time, the connection with the theory of convex polyhedra. Inequalities must have then led him to the then new

[1] In the following historical survey we suppose acquaintance with the notions introduced in the main body of the work.

[2] Fourier gives as the simplest solvability criterion the breaking-off of the reduction.

consideration of half spaces.

8. Gordan and Minkowski. Gordan, who came to the problem fifty years later through invariant theoretic investigations incidental to arithmetic questions, was far removed from such close connections. He stated the elegant transposition theorem in disguised form and proved it in a roundabout way, but then confined himself to diophantine problems, which have their own literature (except for van der Corput) and are remote from the subject treated here.

Minkowski combined the geometric standpoint of Fourier with the n-dimensional treatment of Gordan and handled the question of the solvability of general inequality systems. He discovered the full system of extreme solutions (from which we arrive at the extreme part of the basis) for a single given inequality system and treated the inhomogeneous case as well as the question of dependence of inequalities.

He also introduced many new notions obtained from the geometric point of view: supporting plane, polar body, extreme points, projection space, flat points, corner points (we will not require all these notions) and emphasized the importance of the simplex theorem.

9. Farkas and others. Fourier, Gordan, and Minkowski seem to have worked independently of each other; moreover their results are not well known and must have been discovered many times as new, for example by Steimke and in part by Farkas, Carver, and the author.

Farkas, who was familiar with the literature, attacked the problem many times and displayed his results systematically in 1902. In addition to old results and new extensions to the infinite he deserves proper mention for questions concerning the possible dependences of relations as well as for establishing the general idea of relation. Haar later dealt with the dependence of relations.

In his work on conditionally convergent sequences, Steinitz formulated a general theorem about convex bodies and ray-systems in many-dimensional space which ties in with our considerations. Since then the convex notion has included explicitly neither closure nor boundedness. It is perhaps of interest to note that the notion of supporting line is used constantly by Cramer (see also Newton and de Gua) under the name determinatrice. Stiemke, who follows here chronologically, has been mentioned above.

10. The Americans; concluding remarks. Lovitt solved geometrically a spacial inequality system from political arithmetic (namely from the theory of proportional elections). The participation of Americans who have made attacks on our problem (Dines, Carver, Gummer) begins at this point.

These authors have found interesting isolated results and special solutions. Dines uses elimination for solving inequality systems, but Stokes and Kakeya do not. The method used by Stokes has a somewhat inconvenient geometric interpretation which is dual to ours.

Dines obtains a number from his process which he calls the I-rank. We make no use of this, since, in contrast to the number

of rows of C_1 and C_2 (52) and L_1, L_2, L_3 (59), it is not an affine invariant and thus has only casual meaning.

Carver considers equivalence and unsolvable submatrices, as well as superfluous inequalities in solvable or unsolvable systems. He also derives the relation $r = m-1$ (Theorem D1) for a minimal unsolvable system.

All these investigators have been concerned with a single inequality system, though often in a very general way. We will try to show how their results appear in the general theorems of the theory of inequality systems with variable sign ranges.[1]

11. Bibliography for the main part of the paper. The list is given in chronological order. Numbers refer to the corresponding journals below:

(1) Mathematische Annalen

(2) Mathem und naturwiss. Berichte aus Ungarn

(3) Journal für reine und angewandte Mathematik

(4) Rendiconti del Círcolo Matematico de Palermo

(5) Tôhoku Mathematical Journal

(6) American Mathematical Monthly

(7) Math. és termész. ért.

(8) Annals of Mathematics, 2. series

---

[1] For a fixed choice of sign range, Minkowski first gave a complete solution theory, both for homogeneous and (contrary to an erroneous statement by Stokes) inhomogeneous systems.

(9) Transactions of the American Mathematical Society

(10) Tohoku Imperial Proceedings

(11) Bulletin of the American Mathematical Society

(12) Transactions of the Royal Society of Canada III

(13) Math. es phys. lapok

(14) Mathematische Zeitschrift

(15) Proceedings of the Math. and Phys. Society of Japan, 2. Series

(16) Science Reports Tohoku

1a	1676	Newton	Opuscula (1736)	31-199
1b	1740	de Gua	Usage de l'analyse de Descartes etc.	
1c	1750	Cramer	Analyse des Courbes Algébriques	
1d	1823	Fourier	Oeuvres II	317-328
1e	1829	Gauss	Werke V	23-28
1f	1834	Ostrogradsky	Mém. Ac. Sc. St.-Pétersbourg	
1g	1873	Gordan	(1)	6, 23-28
1h	1885	" - Kerschensteiner	Invariantentheorie, Lpz.	199
1i	1895	Farkas	(2)	12, 263-281
1j	(1898)	Farkas	(2)	15, 25-40
1k	1896	Minkowski	Geometrie der Zahlen, Lpz.1910	39-45
1ℓ	1897	Minkowski	Ges. Abh. II	103-122
1m	(1911)	Minkowski	Ges. Abh. II	131-320
1n	1899	Farkas	(2)	16, 154-157
1o	1901	Farkas	(3)	124, 1-27
1p	1911	Carathéodory	(4)	32, 139-217

1q	1913	Schmidt	(.	74, 271–274
1r	1913	Steinitz	(3)	143, 128–175
1s	1914	Steinitz	(3)	144, 1–40
1t	1914	Farkas	A mechanika alaptanai, Kolozsvár	
1u	1915	Stiemke	(1)	76, 340–342
1v	1915	Kakeya	(5)	8, 218–221
2a	1916	Lovitt	(6)	23, 363–366
2b	1918	Haar	(7)	36, 279–296
2c	1918	Farkas	(7)	36, 297–308
2d	1918	Farkas	(7)	36, 396–408
2e	1918	Dines	(8)	20, 191–199
2f	1921	Carver	(8)	23, 212–220
2g	1924	Haar	Szeged Acta	2, 1–14
2h	1925	Dines	(8)	27, 57–64
2i	1926	Gummer	(6)	33, 488
2j	1926	Dines	(9)	29, 463–470
2k	1926	Dines	(8)	28, 41–42
2l	1927	Dines	(8)	28, 386–392
2m	1928	Fujiwara	(10)	4, 330–333
2n	1928	La Menza	Atti Congr.Int.Mat. Bologna II	199–209
2o	1930	Dines	(11)	36, 393–405
2p	1931	Van der Corput	Proc. Amst.	34, 368–371
2q	1931	Stokes	(9)	33, 782–805
2r	1931	Schlauch	(6)	39, 218–222
2s	1932	Schlauch	(9)	34, 594–619

2t	1932	Stoelinga	Diss. Amsterdam	
2u	1932	La Menza	Bol. Mat.	5, 127-130
2v	1933	La Menza	Bol. Mat.	5, 149-152
2w	1933	Dines-McCoy	(12)	27, 37-70
2x	1934	Bunt	Diss. Amsterdam	98-108
2y	1934	Weyl	Comm. Math. Helvetici	7, 290-306

12. Literature for §§ 15-16.

3a	1637	Descartes	Geometrie Lpz. 1923	73
3b	1828	Gauss	Werke III	65-70
3c	1931	Fourier	Ostw. Klass. 127, Lpz.	92
3d	1834	Jacobi	Werke III	279
3e	1835	Sturm	Ostw. Klass. 143, Lpz. 1904	
3f	1879	Laguerre	Oeuvres I	67-71
3g	1881	Laguerre	Oeuvres I	151-152
3h	1883	Laguerre	Oeuvres I	3-47
3i	1884	Laguerre	Oeuvres I	181-206
3j	1899	Runge	Enz. d. math. Wiss., I, 1	409-416
4a	1912	Fekete	(7)	30, 746-782
4b	1912	Pólya	(7)	30, 783-796
4c	1912	Fekete-Pólya	(4)	34, 89-120
4d	1913	Bálint	(7)	31, 286-305
4e	1916	Bálint	(13)	25, 82-92
4f	1916	Bálint	(13)	25, 178-186
4g	1917	Bálint	(13)	26, 89-106
4h	1917	Curtiss	(8)	19, 251-278

4i	1925	Obreschkoff	Jahresb. d. deutsch Math.-Ver.	33, 52–64
4j	1925	Pólya-Szegö	Aufgaben aus der Analysis, Berlin Bd. 2, V	
4k	1930	Schoenberg	(14)	32, 321–328
4l	1934	Schoenberg	(14)	38, 546–564

13. Further literature. The following books are either cited in the work or were referred to:

5a	Bonnesen-Fenchel	Theorie der donvexen Körper, Berlin 1934
5b	Cesàro	Algebraische Analysis, Deutsche Übertragung von Kowalesdki, 1, Aufl. Lpz. 1904
5c	Hausdorff	Grundzüge der Mengenlehre, Lpz. 1914
5d	Schreier-Sperner	Analytische Geometrie und Algebra I, 1931
5e	van der Waerden	Moderne Algebra I, Berlin 1930

Many of the theorems for finite systems, in particular the transposition theorem, can be generalized to the infinite, but these considerations lie outside the scope of this work. One can assume that the number of unknowns or relations or both is countable or continuous, and obtain many interesting theorems about sequences and integrals. Such theorems, including their applications, and related investigations can be found in:

6a	1913	Kakeya	(5)	3, 137–150
6b	1913	Kakeya	(5)	4, 186–190
6c	1914	Kakeya	(5)	6, 27–31

6d	1914	Kakeya	(5)	6, 130-133
6e	1914	Kakeya	(5)	6, 187
6f	1914	Bohl	(3)	144, 284-313
6g	1915	Kakeya	(5)	8, 14-23
6h	1915	Kakeya	(15)	8, 83-102
6i	1915	Kakeya	(15)	8, 256-261
6j	1915	Kakeya	(15)	8, 403-420
6k	1915	Kakeya	(15)	9, 93-100
6l	1915	Fujiwara	(16)	4, 339-359
6m	1917	Fujiwara	(16)	6, 101-110
6n	1917	Fujiwara	(16)	6, 307-319
6o	1923	Hausdorff	(14)	16, 220-248
6p	1927	Dines	(11)	33, 659-700
6q	1927	Dines	(8)	28, 393-395
6r	1927	Dines	(9)	29, 463-470
6s	1928	Dines	(9)	30, 425-438
6t	1928	Dines	(9)	30, 439-446
6u	1929	Dines	(12)	23, 141-146
6v	1930	Dines	(11)	36, 393-405
6w	1930	McCoy	(11)	36, 878-882
6x	1930	Furtwängler Monatsh. f. Math. u. Phys. (in einer Arbeit von Huber)		37, 59-60
6y	1930	Fujiwara	(10)	6, 297-
6z	1932	Ascoli	Ann. Mat. pura et appl. IV	10, 33-81
6a'	1932	Ascoli	Ann. Mat. pura et appl. IV	10, 203-232
6b'	1932	Schoenberg	(9)	34, 594-619

6c'	1932	Schoenberg	(11)	38, 72-76
6d'	1933	Schoenberg	(11)	39, 273-280
6e'	1933	Schoenberg	(9)	
6f'	1933	Hildebrand -Schoenberg	(8)	34, 317-328

CHAPTER I. GENERALITIES

§3. Introductory definitions and remarks

14. Sets. In general we shall call a set the smallest set with a certain property if none of its proper subsets has that property, while the set itself still has it; in particular this concept is useful when applied to properties which persist when arbitrary elements are added to a set.

If for a given set $\underline{A}$, we have the smallest set $\underline{\bar{A}}$ containing A and having a given property, we shall say that $\underline{\bar{A}}$ is generated by $\underline{A}$ (with respect to the property concerned) and that $\underline{A}$ is a generating set. Every element of $\underline{\bar{A}}$ shall be called dependent on the elements of $\underline{A}$ or generated by $\underline{A}$.

A smallest generating set of a given set $\underline{C}$ shall be called a base of $\underline{C}$. An element of a set, which cannot be generated by the remaining elements, shall be called an independent element. To obtain a clear notion of these concepts consider for instance the property of a set to be linear variety.

"x is an element of $\underline{A}$" we shall write $x \in \underline{A}$, read: x in $\underline{A}$

15. Point sets. In the following we restrict ourselves to the consideration of sets of points of Euclidean n-dimensional space R_n, n = 1,2,···, with a given orthogonal Cartesian coordinate system (the number space). We shall use the notations indicated below:

$\underline{R}_n$ is the whole space. Arbitrary sets of points shall in general be denoted by underlined capitals, finite sets of points

by non-underlined capitals, real numbers by Greek letters. Small Roman letters shall designate single points in $\underline{R}_n$; these points shall also often be identified with the vectors leading to them from the origin.

The origin itself is 0 or 0_n. 1 or 1_n is the unit point of R_n (the points, all of whose coordinates are 1). The unit point of the i-th axis (the vector having on its i-th place a 1 and elsewhere only zeros) shall be denoted by e_i, $i = 1,2,\cdots,n$.

A vector x shall be called ≥ 0, if each of its coordinates $\xi_\nu \geq 0$, $\nu = 1,\cdots,n$; correspondlingly we define > 0, ≤ 0, < 0; in contrast $\neq 0$ shall be the negation of 0.

If $x \geq 0$ or $x \leq 0$, then x will be called <u>definite</u>. For $x > 0$ and $x < 0$ we also write $\mathrm{sg}\, x = +1$ or $\mathrm{sg}\, x = -1$ respectively; x is called <u>properly definite</u>.

16. <u>Matrices</u>. A finite point set A can also be conceived as a rectangular <u>matrix</u>, whose columns consist each of the coordinates of a point a; vice versa evidently every rectangular matrix with n rows and a finite number of columns can be regarded as a finite set of points of R_n.

By interchanging lines with columns we obtain from a matrix A its transposed A'. AB shall mean the matrix product of A and B, (A,B) and $\binom{A}{B}$ the matrix obtained from A and B respectively by juxtaposition and by writing one below the other. r shall in general designate the rank of A, n the number of rows, and m the number of columns. In particular E_n shall be the unit matrix

of order n (in the main diagonal everywhere 1, elsewhere everywhere 0). For square A let $|A|$ be the value of the determinant of A.

When forming submatrices from A by omission of certain rows and columr we always suppose the order of the remaining rows and columns unchanged. In particular the value of subdeterminants of A is to be determined accordingly. Each submatrix is the intersection of a combination of rows with a combination of columns.

By A_i we shall mean a submatrix of the rectangular matrix A which has rank i and consists of i columns of A. Here $1 \leq i \leq r$. However, the notation A_i is defined in this way only for the letter A.

17. Equations. We recall some facts from the theory of linear equations which we shall need. Inhomogeneous systems are solvable exactly in the case where the matrix of the coefficients has the same rank as the matrix of the coefficients and the free terms, and one obtains their solutions from quotients of subdeterminants.

Solutions of a homogeneous system $xA = 0$ of rand r with $n > r$ unknowns can be combined linearly and homogeneously from $n-r$ from among them. A matrix consisting of $n-r$ such solutions as its rows shall be denoted as an (equational) base G or G_A. If M is a variable square matrix of order $n-r$, whose determinant does not vanish, then MG represents every other G_A (transformation of the base).

Every A_i, $i < r$, corresponds to a subsystem of the entire system of equations. Let G_1 be a G of A and G_2 a G of A_i.

The rows of G_1 are solutions of the whole A-system and hence a priori of the A_1-system, but do not form a complete system of solutions since their number is by $r-1$ too small. The missing $r-1$ solutions can be taken from the G_2, as an arbitraty combination of $r-1$ rows which are independent from G_1 and from each other. We thus have the often employed

Theorem A1.[1] Every G_A can be extended to a G of A_1 by including $r-1$ rows out of a given G_{A_1}.

18. Solution by determinants. Incidentally a G_A can be obtained in the following way (see for instance, Cesaro, Algebraische Analysis, page 47).

A fixed A_r contains $\binom{n}{r+1}$ submatrices $A_{r',r+1}$ consisting of r columns and $r+1$ rows. Now consider one such $A_{r',r+1}$. To each of its rows, we let correspond the value of the determinant of order r formed by the other rows, after multiplying these $r+1$ determinants alternatingly by 1 and -1. To the other $n-r-1$ rows of our A_r we assign a 0 and thus obtain from every $A_{r',r+1}$ n numbers corresponding to the rows of A_r (that is, the rows of A); these n numbers we write in unchanged order as a row. Altogether we obtain thus $\binom{n}{r+1}$ rows, among which there are always at least $n-r$ independent ones; any $n-r$ independent rows of the matrix thus obtained form a G_A. Hence we have:

---

[1] "Steinitz' Exchange Theorem"; see for instance, B. Schreier-Sperner, Analytische Geometrie, and van der Waerden, Moderne Algebra, page 96.

Theorem A2. A G_A can be formed from subdeterminants of order r of A and zeros.

19. Linear sets. We consider as known the main results of the theory of linear manifolds (varieties, spaces, sets) and mention them only briefly.

The linear manifold through the origin $\underline{O}_n$ generated by m independent points ("spanned by m vectors") is called of dimension m, likewise every parallel manifold.

Two linear manifolds of dimensions m_1 and m_2 whose intersection exists and has dimension d, generate together a manifold of dimension m_1+m_2-d (for the generation choose d points from the intersection, m_1-d from the first, and m_2-d from the second manifold). If $m_1+m_2 = n$ and $d = 0$, then the two varieties are called complementary (with respect to $\underline{R}_n$).

Hyperplane means a linear variety of dimension $n-1$, plane always one of dimension two.

Every linear equation $xa = 0$ with $a \neq 0$ represents a hyperplane orthogonal to the vector a. More generally: the system of equations $xA = 0$ is solved by the points of a linear variety which is complementary to the linear variety generated by A and orthogonal to it at the origin. Theorem A1 can then be interpreted by saying that of these two varieties one increases when the other decreases.

Finitely many linear manifolds are called linearly independent if none of them lies in the linear manifold which is generated by the others together.

§4. Convex sets.

20. Definitions. As we shall see the convex sets correspond to the linear systems of inequalities in the same sense in which linear manifolds belong to systems of equations. A set is called convex if, together with any two of its points x aand y, it contains the whole segment $\lambda x+(1-\lambda)y$, $0 \leq \lambda \leq 1$ connecting them. It is true that the most general convex sets belong to infinite systems of inequalities.

A convex set which contains no straight line shall be called linefree.

By a convex polyhedron (polytope) we mean a convex set which can be generated by finitely many points x_i, that is, the set of all $\Sigma\lambda_i x_i$ with $\Sigma\lambda_i = 1$. In particular $n+1$ linearly independent points determine a simplex (generalization of the triangle in R_2 and the tetrahedron in R_3).

In R_1 the concept of polyhedron degenerates into that of a segment or a single point. In addition there exists in R_1 no further convex set except the whole R_1 and the ray (half line). The ray and segment can also be open, the segment also half open. Thus all connected sets are here convex while otherwise only the converse holds.

21. The boundary. A convex set need not be closed. The closure of a convex set is always convex (as seen very easily although I do not find it mentioned explicitly). Polyhedra are closed, linear manifolds are convex and closed.

If the closure of a convex set A has points in common with

a hyperplane but if all other points of $\underline{A}$ lie only on one side of the hyperplane, then as usual the hyperplane shall be called <u>support</u> of A.

A convex set which, together with each of its points x and each of its boundary points y, contains the whole segment from x to y, except at most the point y, shall be called <u>wholefaced</u> (it contains every open face (=linear piece of boundary) either entirely or not at all). To the wholefaced sets belong the closed convex sets and the one-dimensional convex sets.

22. <u>Intersection</u>. We recall the simple fact that the property to be convex, like that to be linear, bounded or closed, remains unchanged by intersection.

Now let arbitrarily many wholefaced sets $\underline{A}$ be given, and let their intersection exist and be called $\underline{D}$. Let $\underline{\bar{A}}$ be the closure of $\underline{A}$, and $\underline{\bar{D}}$ the intersection of all $\underline{\bar{A}}$. Then if x is a point of $\underline{D}$, y a point of $\underline{\bar{D}}$, we also have $x \in \underline{A}$, $y \in \underline{\bar{A}}$, for all $\underline{A}$. Because of the wholefacedness of $\underline{A}$ also every point of the open segment from x to y lies in every $\underline{A}$, therefore also in $\underline{D}$. Hence $\underline{\bar{D}}$ is the closure of $\underline{D}$ and we have:

<u>Theorem F1</u>. <u>The intersection of wholeface sets is itself wholefaced.</u>

At the same time we obtain:

<u>Theorem F2</u>. <u>For wholefaced sets the formation of closure and intersection commute, if the intersection exists.</u>

It is now easily seen that the projection of a wholefaced set on a linear variety is also wholefaced.

23. Convex cones. If a convex set $\underline{A}$ can be regarded as the union of open rays with fixed end-point a, possibly together with the point a itself, then $\underline{A}$ shall be called a convex cone (ray domain) with vertex a. Together with x the set A contains also the ray $\overline{ax}$, that is, the set of all $a+\lambda(x-a)$, $\lambda > 0$.

If $a = 0$ then $\underline{A}$ contains with x also λx, $\lambda > 0$, and with x and y also $\frac{x+y}{2}$ (as a convex set), hence also $2 \cdot \frac{x+y}{2} = x+y$. **Therefore $\underline{A}$ contains in this case also together with finitely many points every homogeneous linear combination of these points with positive coefficients as a sum of terms $\lambda_i x_i$.** (Incidentally this property could also have served as a definition of convex cones with vertex 0, since it immediately entails convexity: $\lambda x+(1-\lambda)y$ being a special linear combination.)

The rays leading from a given point a to all points of a convex set $\underline{A}$ form the convex cone with vertex a which is generated by $\underline{A}$ (or also by every generating set of $\underline{A}$), if a is included exactly in case little a belongs to $\underline{A}$; this cone we denote in the sequel always by $\overline{a\underline{A}}$. This notation includes as a particular case that of the ray $\overline{ax}$.

In $\underline{R}_1$ a set $\overline{a\underline{A}}$ can be the whole $\underline{R}_1$, a closed or open ray or a single point.

A convex cone containing its vertex and generated by finitely many points shall be called polyhedral.

24. Wholefacedness. If several wholefaced convex sets $\underline{A}_i$, $i = 1, \cdots, m$, are given such that a belongs to the closure of their intersection $\underline{A}$, that is, belongs to all of them or to some and lies on the boundary of the other, then we shall see easily

that $\overline{aA}$ *is also the intersection of the* $\overline{aA_i}$, which need not hold in the general case (if a lies elsewhere).

For if a point p belongs to $\overline{aA}$ then it certainly belongs to all $\overline{aA_i}$; and conversely if p belongs to every $\overline{aA_i}$, $i = 1,\cdots,m$, then the ray $\overline{ap}$ contains a point p_i of each A_i, and that point p_1 which lies nearest to a must also (because of the wholefacedness of the sets A_i) belong to all A_i, that is, to A, whence $p \in \overline{aA}$. Here we supposed $p \neq a$; for $p = a$ the contention is trivial. We have thus shown:

Theorem F3. *For finitely many wholefaced sets the formation of convex cone and intersection commute, if the vertex belongs to the closure of the intersection.*

Furthermore it is easy to convince oneself that $\overline{aA}$ for $a \in A$ and wholefaced A only depends on the part of A which is situated in an arbitrarily small neighborhood of a.

25. *The number line.* The one-dimensional space R_1 shall also be denoted by E_{-0+}, the set of all positive numbers by E_+, the set of non-positive numbers by E_{-0} etc. We have six convex sets E, namely E_{-0+}, E_{-0}, E_{0+}, E_-, E_+, E_0 (zero only).

Then evidently the product of each of these six sets with a fixed number is again identical with one of these sets, likewise the sum of two and therefore of finitely many from among these six sets is again one of these sets. Thus if each variable ξ_i of a linear form $\Sigma\lambda_i\xi_i$ ranges over a given E_i we see that the range of values of the linear form is also such an E.

The six $\underline{E}$ shall be called one-dimensional **sign ranges**. Four of them: $\underline{E}_{-0+}$, $\underline{E}_{-0}$, $\underline{E}_{0+}$, and $\underline{E}_{0}$, are closed.

26. Sign ranges. Let $\underline{E}_1, \cdots, \underline{E}_n$ be n one-dimensional sign ranges. Then the set of points of $\underline{R}_n$ whose first ,···, n-th coordinate is respectively in $\underline{E}_1, \cdots, \underline{E}_n$, shall be called an n-dimensional sign range $\underline{V}$, more explicitly $\underline{E}_1 \cdot \underline{E}_2 \cdot \ldots \cdot \underline{E}_n$. Every $\underline{V}$ is a convex cone with vertex O_n, contains if closed $\underline{E}_0 = \underline{E}_0 \cdot \ldots \cdot \underline{E}_0$ (n times) $= O_n$ and is contained in $\underline{E}_{-0+}$ (that is, the whole space).

In the same way in which $\underline{R}_2$ is divided in quadrants and $\underline{R}_3$ in octants, $\underline{R}_n$ is decomposed by the n coordinate hyperplanes in 2^n parts which may be called orthants because of their orthogonal angle. They are special sign ranges, namely those in whose formation neither $\underline{E}_0$ nor $\underline{E}_{-0+}$ takes parts. $\underline{V}_{0+} = \underline{E}_{0+}{}^n$ shall be called the principal orthant. It is closed, and its interior is the orthant $\underline{V}_+ = \underline{E}_+{}^n$.

CHAPTER II. REPRESENTATION OF SOLUTIONS

§5. Inequalities

27. Relations. In what follows, we shall consider only equations and inequalities and shall include both under the term relations.

A single relation in n real unknowns will be written in the form

$$P_{a,\gamma,E}: \qquad xa + \gamma = \omega, \qquad \omega \in E$$

or simply $xa+\gamma \in \underline{E}$, where, as later in the same context, x is a row of n unknowns and a is a coefficient column, γ is a constant and $\underline{E}$ is one of the six one-dimensional sign ranges.

$\underline{E}_0$, $\underline{E}_{-0+}$, $\underline{E}_-$, $\underline{E}_+$, $\underline{E}_{-0}$, $\underline{E}_{0+}$ correspond respectively to the signs $=0$, $\gtrless 0$, <0, >0, ≤ 0, ≥ 0. We distinguish between equations, identities, and inequalities, and indeed call the $\gtrless$-relation, as well as the relation with $a = \gamma = 0$ identities (since they are satisfied for all x), and the $>$ and $<$ relations strict inequalities. Relations with $\gamma = 0$ are homogeneous.

The points satisfying an equation constitute a hyperplane. Those satisfying a corresponding inequality make up an open or closed half space defined by the hyperplane. An identity represents the whole space $\underline{R}_n$. All these point sets are wholefaced.

28. Systems. A set of m such relations P forms a finite relation system, which can obviously be written in the form

I $\qquad \Sigma_{A,c,\underline{V}}: \qquad xA+c \in \underline{V},$

where A is a coefficient matrix, c is a column of constants, and the m-dimensional sign range $\underline{V}$ is uniquely determined by the given sign combination. Each column a of A corresponds to a relation P of Σ.

The totality of solutions of the system Σ (when such exist) will be called the <u>solution domain</u> and denoted by $\underline{L} = \underline{L}(\Sigma)$, so that the most general solution of I has the form

I': $\qquad x \in \underline{L}.$

Since $\underline{L}$ is an intersection of half spaces and hyperplanes, it is a wholefaced set. In the case of homogeneous relations (c=0), $\underline{L}$ is a convex cone with vertex 0.

Not all wholefaced sets or convex cones, however, can be obtained as the $\underline{L}$ of certain Σ's. Rather it happens only for special sets, which for reasons which will be evident later, we shall call <u>polyhedral</u> (unrund). It follows from the definition that the intersection of polyhedral sets is polyhedral.

29. <u>Remarks</u>. Evidently $\underline{V}$ is wholefaced and polyhedral.

Systems of equations correspond to the special case $\underline{V} = \underline{E}_0{}^m$; more generally, the case E^m is of particular interest. All others can be built up from systems of this last type through simple logical and arithmetic operations.

The addition of identities to Σ means an enlargement of A by certain columns, while $\underline{V}$ is multiplied by a product of factors

of the form $\underline{E}_{-0+}$ or $\underline{E}_0$. Of course $\underline{L}$ is unchanged.

We note that the task of finding those $x \in \underline{L}$ which are in a second given sign range $\underline{V}_1$, is nothing new; indeed one has but to add special homogeneous relations for x, and then the complete system can be written in the form

$$x(E,A)+\binom{0}{c} \in \underline{V}_1\underline{V}.$$

3C. Subsystems. Corresponding to each point $x \in \underline{R}_n$ we can associate a subsystem $\Sigma^{(x)}$ of Σ as follows: $\Sigma^{(x)}$ shall contain all equations of Σ but only those inequalities in which x produces an equality sign (i.e., x lies in the associated hyperplane); this subsystem will be called the x-corner[1] of Σ.

Let $\overline{\underline{V}}$ be the closed sign range corresponding to $\underline{V}$, and denote by $\overline{\underline{L}}$ the solution domain of

$$\overline{\Sigma}: \qquad xA+c \in \overline{\underline{V}}.$$

We suppose $\underline{L}$ is non-void, so that by Theorem F2 the closure of $\underline{L}$ is $\overline{\underline{L}}$. We then consider the convex cone $x\underline{L}$ generated by $\underline{L}$ with vertex x, where $x \in \overline{\underline{L}}$, first of all in the case of a single relation P.

If the $\underline{L}$ of P is either a hyperplane or the whole space, then $\overline{x\underline{L}} = \underline{L}$, and similarly if $\underline{L}$ is a halfspace and x a boundary point of $\underline{L}$; if x is an interior point of the halfspace $\underline{L}$, then $\overline{x\underline{L}} = \underline{R}_n$. These cases can be brought together in the formula $\overline{x\underline{L}(P)} = \underline{L}(P^{(x)})$, if we suppose the solution domain of an empty system is $\underline{R}_n$.

---

[1] See the Minkowski "projection space."

In the general case of a system Σ of many relations, by Theorem F3 $\overline{x\underline{L}(\Sigma)}$ is the intersection of all m single $\overline{x\underline{L}(P)}$, where the $\underline{L}(P)$ are hyperplanes, halfspaces, or $\underline{R}_n$. Consequently, $\overline{x\underline{L}}$, for $x \in \underline{L}$, is the set $\underline{L}(\Sigma^{(x)})$ of solutions of $\Sigma^{(x)}$:

<u>Theorem</u> E1. <u>If</u> $x \in \underline{L}$, <u>then</u> $\overline{x\underline{L}} = \underline{L}(\Sigma^{(x)})$.

31. <u>The coefficient range of a row.</u> We restrict our attention to homogeneous systems (up to §9 inclusive), which we think of as given in the form

$$\text{II} \qquad \Sigma_{A,O,\underline{V}}: \qquad xA \in \underline{V}.$$

Let ℓ be a point of $\underline{R}_n$. We can form all of its proportional points $\omega\ell$, where ω is an arbitrary number, and ask ourselves for which values of ω the point $\omega\ell$ is a solution of II, i.e., lies in $\underline{L}(\Sigma)$. Now $\underline{L}$ contains with each point all its positive multiples, so only six cases can arise: ℓ, $-\ell$, and 0 in $\underline{L}$; ℓ and 0 in $\underline{L}$, $-\ell$ not in $\underline{L}$; ℓ in $\underline{L}$, 0 and $-\ell$ not in $\underline{L}$; $-\ell$ and 0 in $\underline{L}$, ℓ not in $\underline{L}$; $-\ell$ in $\underline{L}$, ℓ and 0 not in $\underline{L}$; ℓ and $-\ell$ not in $\underline{L}$, 0 in $\underline{L}$; in these cases the values of ω form respectively the one dimensional sign ranges $\underline{E}_{-o+}$, $\underline{E}_{o+}$, $\underline{E}_{+}$, $\underline{E}_{-o}$, $\underline{E}_{-}$, and $\underline{E}_{o}$. The associated $\underline{E}$ will be called the <u>coefficient range</u> $\underline{E}_\ell$ <u>belonging to the row</u> ℓ. If $\underline{V}$ is closed, so is $\underline{E}_\ell$.

32. <u>The coefficient range of a matrix.</u> If L is a matrix of n columns and p rows $\ell_1, \cdots, \ell_p$, then a linear combination $\Sigma\omega_i\ell_i$ of these rows lies in $\underline{L}$ if (but not in general only if) each ω_i is in the corresponding coefficient range $\underline{E}_{\ell_i}$. Hence if we denote

the row of ω_i by w and the p-dimensional sign range $\underline{E}_{\ell_1}\cdots\underline{E}_{\ell_p}$ by $\underline{W}$, then $w \in \underline{W}$ implies $wL \in \underline{L}$. $\underline{W}$ is closed with $\underline{V}$.

Let $\underline{W}L$ denote the set of all wL for $w \in \underline{W}$. Then $\underline{W}L$ is a subset of $\underline{L}$, but $\underline{W}$ may not be the largest set, indeed not even the largest sign range, having this property; however, for closed $\underline{V}$ and $\underline{W}$ the latter is valid: for then the $\omega_i \ell_i$ themselves belong to $\underline{W}L$ (since for all $k \neq i$, ω_k can be taken as 0), so that $\underline{E}_{\ell_i} \ell_i$ is a subset of $\underline{L}$, where $\underline{E}_{\ell_i}$ is the i-th coordinate range of $\underline{W}$.

$\underline{W} = \underline{W}(\Sigma,L)$ will later play an important role: we call $\underline{W}$ the <u>coefficient range belonging to the matrix</u> L. The coefficient range of a matrix is thus the product of the coefficient ranges of the rows of the matrix.

Since $\underline{W}L$—as well as $\underline{W}$—contains with two points their sum and with each point any positive multiple of it, $\underline{W}L$ is a convex cone with vertex 0_n.

33. <u>Solution matrices</u>. It is clear that $\underline{W}L$ represents solutions, but it may not represent all solutions; on the other hand, it is conceivable that $\underline{W}L = \underline{L}$, so that each solution x of II can be written in the parametric form

$$\text{II}' \qquad x = wL, \quad w \in \underline{W}.$$

If this is the case for each closed $\underline{V}$, i.e., for each combination of relation signs (except > and <), then $\underline{W}(\underline{V})L = \underline{L}(\underline{V})$ identically in $\underline{V}$, and we call L a <u>solution matrix</u> of A.

If L is a solution matrix, it remains one after the addition of arbitrary rows. For one can represent a point of $\underline{L}$ with the new matrix by using the old representation, supplying zero coefficients for the new rows; or it may be that essentially different, new representations can be obtained.

Accordingly, we will look for those L which contain no superfluous rows; they might be called <u>bases of</u> A. As one of our principal results it will be shown in the next paragraph that a basis of A exists for each A (Theorem C2). A matrix L is a solution matrix of A if and only if it contains a basis of A.

If we have a relation system with A_1, a subset of the columns of A, as coefficient matrix, we can interpret it as a special system with A as coefficient matrix, the completion to A being effected by adding identities. Hence the A_1-systems are subsystems of the A system and each solution matrix L_A of A is a solution matrix L_{A_1} of A_1.

§6. <u>Existence of</u> L.

34. A^*. In this section A^* is a matrix of n rows and $n+1$ columns, whose first n columns form the identity matrix E_n and whose last column a consists of n arbitrary, not all zero, numbers $\alpha_1,\cdots,\alpha_n$: $A^* = (E_n,a)$.

For each pair h,k, $1 \leq h < k \leq n$, we form a vector ℓ_{hk}, whose h-th coordinate is α_k and k-th coordinate is $-\alpha_h$, the others being zero: $\ell_{hk} = \alpha_k e_h - \alpha_h e_k$. Since $\ell_{hk} a = 0$ and ℓ_{hk} has at most two non-zero coordinates, the ℓ_{hk} are points on the intersection of the hyperplane $xa = 0$ with the $\binom{n}{2}$ two-dimensional coordinate planes.

We write these $\binom{n}{2}$ rows ℓ_{hk} in lexicographic order as to h and k and obtain a matrix L_1 depending on a, $L_1 = L_1(a)$. Now we adjoin to this the n rows e_i, getting a matrix $L = \binom{L_1}{E_n}$ with $\binom{n+1}{2}$ rows. (In the trivial case n=1, we take $L = E_1$). We shall prove:

<u>Theorem C1. L is a solution matrix of</u> A^*.

The proof of this assertion is presented in 35–37.

35. <u>Proof.</u> If x is any solution of the closed homogeneous relation system $xA^* \in \underline{V}$, we are to prove that x has the form wL, $w \in \underline{W}$. Let $x = \Sigma\xi_i e_i$. Then our system has the form

$$\text{III} \qquad \Sigma A^*: \qquad \xi_i \in \underline{E}_i, \quad i = 1,\cdots,n; \qquad \sum_{i=1}^{n} \alpha_i \xi_i \in \underline{E}_{n+1}$$

(the first n relations belong to the columns of E_n, the last to the column a).

Now in addition to x either all vextors $\xi_i e_i$ belong to $\underline{L}(\Sigma)$, or not all (perhaps none) do. Two cases therefore arise, the first of which is handled immediately, the second in 36–37.

Case 1: Each $\xi_i e_i$ is a solution of $xA^* \in \underline{V}$. Since e_i is a row of L, and $\underline{W}L$ by 32 contains all $\omega_i e_i$ which are in $\underline{L}$, the $\xi_i e_i$ are in $\underline{W}L$, hence so is their sum.

36. <u>Case</u> 2. Suppose some $\xi_h e_h$ does not belong to $\underline{L}$. If we insert the components of $\xi_h e_n$ in system III, we see that the first n relations are fulfilled (the value of the linear forms on the left in III remain the same as for x or vanishes). Consequently the $(n+1)^{st}$ relation $\alpha_h\xi_h \in \underline{E}_{n+1}$ does not hold, whence

(since we have a closed system) $\alpha_h\xi_h \neq 0$. If all the remaining non-zero $\alpha_i\xi_i$ had the same sign as $\alpha_h\xi_h$, then their sum ax could not lie in $\underline{E}_{n+1}$, contrary to assumption. Thus at least one $\alpha_k\xi_k$ has sign opposite that of $\alpha_h\xi_h$, and we pick one such.

Since $\xi_h\alpha_h\xi_k\alpha_k < 0$, then $\xi_k \in \underline{E}_k$ implies $\frac{-\xi_h\alpha_h}{\alpha_k} \in \underline{E}_k$ and $\xi_h \in \underline{E}_h$ implies $\frac{-\xi_k\alpha_k}{\alpha_h} \in \underline{E}_h$. This, together with $\ell_{hk}a = 0$, shows that $\frac{\xi_h}{\alpha_k}\ell_{hk}$ and $\frac{-\xi_k}{\alpha_h}\ell_{hk}$ are in $\underline{L}$; but since ℓ_{hk} is a row of L, $\frac{\xi_h}{\alpha_k}\ell_{hk}$ and $\frac{-\xi_k}{\alpha_h}\ell_{hk}$ are thus in $\underline{WL}$.

37. The reduction. Now set

$$x_1 = x - \frac{\xi_h}{\alpha_k}\ell_{hk}, \qquad x_2 = x + \frac{\xi_k}{\alpha_h}\ell_{hk}$$

$$x = x_1 + \frac{\xi_h}{\alpha_k}\ell_{hk} = x_2 - \frac{\xi_k}{\alpha_h}\ell_{hk};$$

then x differs from x_1 and x_2 only in the h^{th} and k^{th} components, indeed ξ_h and ξ_k, the components of x, are different from zero, as we saw, while x_1 has a zero in the h^{th} position, x_2 a zero in the k^{th} position.

Thus, since $\ell_{hk}a = 0$, also

$$x_1a = x_2a = xa \in \underline{E}_{n+1}$$

$$x_1e_i = x_2e_i = xe_i \in \underline{E}_i, \quad i = 1,\cdots,n, \quad i \neq h,k.$$

Moreover, $x_1e_h = x_2e_k = 0$. Hence x_1 satisfies all relations of Σ except the k^{th}; this one says that the k^{th} component of x_1 is

in $\underline{E}_k$,

$$\xi_k + \frac{\xi_h \alpha_h}{\alpha_k} \in \underline{E}_k.$$

Similarly the condition that $x_2 \in \underline{L}$ is

$$\xi_h + \frac{\xi_k \alpha_k}{\alpha_h} \in \underline{E}_h.$$

If $|\xi_h \alpha_h| \leq |\xi_k \alpha_k|$, then $|\xi_k| \geq |\frac{\xi_h \alpha_h}{\alpha_k}|$, so that $\xi_k + \frac{\xi_h \alpha_h}{\alpha_k}$ has the same sign as ξ_k or is zero. Since

$$\xi_k \in \underline{E}_k, \quad \xi_k \neq 0, \quad 0 \in \underline{E}_k,$$

$\xi_k + \frac{\xi_h \alpha_h}{\alpha_k} \in \underline{E}_k$ and $x_1 \in \underline{L}$. Similarly $|\xi_k \alpha_k| \leq |\xi_h \alpha_h|$ implies that $x_2 \in \underline{L}$.

Therefore at least one of the points x_1, x_2 is in $\underline{L}$, and if we denote this one by $\bar{x}$, we have split x into two parts, of which one, $x-\bar{x}$, is in $\underline{WL}$, as we saw above, while $\bar{x}$ is in $\underline{L}$ and has at least one zero more than x. Repetitions of this process must terminate after no more than n-1 steps, and we come finally to zero or back to case 1.

This completes the proof that L is a solution matrix of A^*.

38. <u>General</u> A. We shall prove

<u>Theorem C2. Every matrix</u> A <u>has a solution matrix.</u>

The proof is by induction on m, the number of columns of A. For m=1 we have the case just studied, if we recall that a solution matrix of $A^* = (E_n, A)$ is a solution matrix of A (33).

Now let L be a solution matrix of an arbitrary A. We look

for a solution matrix of the matrix (A,a) with the column a adjoined. This adds to II another relation: $xa \in \underline{E}_{m+1}$. By the induction hypothesis we can replace the first m relations by their solution $x=wL$, $w \in \underline{W}(L)$, so we have

$$wLa \in \underline{E}_{m+1}, \quad w \in \underline{W}$$

or (for suitable E)

$$w(E,La) \in \underline{W} \cdot \underline{E}_{m+1}$$

Since La is a column, this is a relation system in w of the special type handled in 37. Hence (Theorem C2 and 33) it is solved by

$$w = u\binom{L_1(La)}{E}, \quad u \in \underline{U}$$

for a suitable sign range $\underline{U}$. Thus for x we have the general form

$$x = u\binom{L_1(La)}{E}L = uL_2, \quad u \in \underline{U},$$

in which $L_2 = \binom{L_1(La)}{E}$ is a solution matrix of (A,a).

§7. Specification of all L.

39. An explicit L. Without using induction or this simple existence proof, we can write down explicitly for each A a solution matrix L_A.

If we consider those special $\underline{V}$ composed of only E_0 and E_{-0+}, then after omitting identities, the relation system II becomes an equation system Σ belonging to a certain subset $\bar{A}$ of

the columns of A. Since for equations, ℓ is a solution if and only if $-\ell$ is, only E_0 and E_{-0+} occur in $\underline{W}(\Sigma)$ and we see that a necessary and sufficient condition for $\overline{W}(\Sigma)L = L(\Sigma)$ (i.e., for the representation of each solution in terms of solutions appearing among the rows of L) is that L contain a solution basis $G_{\overline{A}}$ of $\overline{A}$.

Strangely, the fulfillment of the conditions imposed by all these special $\underline{V}$ suffices for the remaining $\underline{V}$. In other words, the following theorem, whose proof is the aim of this section, is valid.

<u>Theorem C3.</u> <u>A matrix</u> L, <u>which for each</u> A_1 <u>of</u> A <u>contains all rows of a</u> G_{A_1}, <u>is a solution matrix of</u> A.

Since there are only finitely many A_1, it is only necessary to form a G_{A_1} for each, and write these together to get an L. The case $\underline{V} = \underline{E}_{-0+}{}^m = \underline{R}_m$ should not be overlooked; here any non-singular n by n matrix can be selected: this basis we denote by G_{A_0}, A_0 arising from A so to speak by the omission of all its columns.

40. <u>Proof.</u> If x is a solution of system II, we will show that x can be represented as $\Sigma\omega_1\ell_1$ and is thus in $\underline{W}L$. We recall the definition of the x-corner $\Sigma^{(x)}$ of Σ: $\Sigma^{(x)}$ contains all equations of Σ and those inequalities P of Σ such that $P(x) = 0$. For each column a of the coefficient matrix $A^{(x)}$ of $\Sigma^{(x)}$ we have $xa = 0$. Thus x can be represented as a linear combination of the rows g of $G_A(x)$. Since $G_A(x)$ occurs

in L, the g are certainly among the l_i.

If we consider now those columns $a_1, a_2, \cdots$ of A, which first of all do not appear in $A^{(x)}$ and secondly belong to an inequality of Σ (not to an identity), we can distinguish three cases:

(1) there are no such columns;

(2) there is an a_1 such that for each other a_i a relation of the form $a_i = \lambda_1 a_1 + \Sigma\lambda a$, $\lambda_1 \neq 0$, exists, where the a without indices are columns of $A^{(x)}$. In other words, there is just one inequality modulo $A^{(x)}$;

(3) more than one inequality modulo $A^{(x)}$ exists.

The first two cases will be handled in 41, the third case afterward.

41. First and second cases. In the first case, where only identities appear outside of $A^{(x)}$, the representation of x as an arbitrary linear combination of the g is already of the form $\Sigma \gamma_i \ell_i$, for both g and $-g$ are in $\underline{L}$, so the coefficient range of g is $\underline{E}_{-0+}$.

In the second case $xa_1 \neq 0$, which means there is at least one $g = g_0$ with $g_0 a_1 \neq 0$. Each a_i, $i>1$, has by assumption the form $\lambda_1 a_1 + \Sigma\lambda a$, whence

$$g_0 a_i = \lambda_1 g_0 a_1 + \Sigma\lambda g_0 a = \lambda g_0 a_1$$

and $xa_i = \lambda_1 x a_1$. Let $\frac{xa_1}{g_0 a_1} = \mu$. Then also $\frac{xa_i}{g_0 a_i} = \mu \neq 0$, or $\mu g_0 a_i = xa_i$. Thus $\mu g_0 \in \underline{L}$.

If in addition to the relation system $\Sigma^{(x)}$ belonging to

$A^{(x)}$, we consider the one belonging to $(A^{(x)}, a_1)$, we know from Theorem A1 that a basis of $G_A(x)$ can be obtained by adding to the rows g_1 of $G_{(A(x), a_1)}$ some row independent of them, so that x is representable as a linear combination of the g_1 and this rown. But we can choose g_o as such a row, as $g_o a_1 \neq 0$ implies that g_o is independent of the g_1. Since $g_1 a_1 = \lambda_1 g_1 a_1 + \Sigma\lambda g_1 a = 0$, both g_1 and $-g_1$ are in $\underline{L}$.

The coefficient of g_o in the representation of x must have the same sign as μ, since μg_o gives the inequalities the same sign as x. Thus the desired form for x is accomplished and the proof in this case is complete.

42. <u>The third case</u>. We retain the notation of the preceding paragraph.

In this case there is an a_2 independent of $(A^{(x)}, a_1)$; consequently there is a g_1 with $g_1 a_2 \neq 0$. We form the set of points $\xi x + \lambda g_1$, where ξ and λ are arbitrary numbers.

Which of the points $\xi x + \lambda g_1$ are in $\underline{L}$, i.e., satisfy $\xi x A + \lambda g_1 A \in \underline{V}$?

This is a system of m homogeneous relations in two unknowns. However, because $xa = g_1 a = 0$, the relations corresponding to $\Sigma^{(x)}$ and the identities are fulfilled for all ξ and λ; thus only inequalities of the form $\xi x a_i + \lambda g_1 a_i \geq 0$, resp. ≤ 0, $i \geq 1$, remain, and these can be written as

IV $$\xi \geq -\lambda \frac{g_1 a_1}{x a_1} ,$$

since the sign $\leq$, in conjunction with the fact that $\xi = 1$, $\lambda = 0$

is a solution, gives a contradiction. As in each system of inequalities of the form $\xi \geq \gamma\alpha_1$, the system IV can be replaced by those two inequalities of the system corresponding to α', the smallest, and α'', the largest, of the α_1. Since $\frac{g_1 a_1}{xa_1} = 0$, $\frac{g_1 a_2}{xa_2} \neq 0$, not all the α_1 are equal, and $\alpha' < \alpha''$.

43. The reduction. Since $\xi = \alpha''$, $\gamma = 1$ inserted in IV gives $\alpha'' \geq \alpha_1$, and $\xi = -\alpha'$, $\gamma = -1$ gives $-\alpha' \geq -\alpha_1$, both are solutions of IV. Therefore

$$x_1 = -\alpha' x - g_1$$

$$x_2 = \alpha'' x + g_1$$

are in L, and $x = \frac{x_1 + x_2}{\alpha'' - \alpha'}$ is a positive homogeneous form in x_1 and x_2.

But now $A^{(x_1)}$ contains at least one more column than $A^{(x)}$, since $x_1 a = 0$ and $\xi = -\alpha'$, $\gamma = -1$ produces an equality sign in one of the inequalities IV, which can happen only from an equality sign in II; similarly $A^{(x_2)}$ contains at least one more column than $A^{(x)}$.

If we decompose x_1 and x_2, if possible, in the same way, the process must terminate after a finite number of steps. We will then have x as a combination of solutions to each of which the process is no longer applicable and accordingly we are back in case 1 or case 2. Thus these solutions are in $\underline{W}L$, and since $\underline{W}L$ is a convex cone with vertex zero, x, being a positive homogeneous combination of elements of $\underline{W}L$, is also in $\underline{W}L$.

This completes the proof that L is a solution matrix of A.

§8. The solution basis.

44. Preparation. We consider a submatrix A_i (that is, one consisting of i linearly independent columns of A) and an A_{i-2} formed from A_i by omitting two columns. If we omit only one or the other of the two columns, we obtain in each case an A_{i-1}. For each of these A_{i-1} there is a G, as was noted earlier, which consists of the rows of a given G of A_i and an additional independent row.

Since both A_{i-1} are independent of each other and therefore generate distinct linear systems, the G of both A_{i-1} are independent of each other.

If now we add to the n-i rows of the given G_{A_i} both of the rows which we have used singly in the representation of the $G_{A_{i-1}}$, we have n-i+2 independent rows, which are orthogonal to the columns of A_{i-2} and thus form a G of A_{i-2}. This reasoning is also valid for i=2.

Thus we see that a matrix which contains a G of each A_i as well as a G of each A_{i-1}, also contains a G of each A_{i-2}. Through recursion on i we thus get

Theorem C4. Each matrix which contains a G of each A_r and A_{r-1}, where r is the rank of A, is a solution matrix of A.

45. Construction. Now all A_r of A generate the same linear system, and the G of these A_r are identical with that of A. If we call this G_A, then we can produce each $G_{A_{r-1}}$ by adding a suitable row to the n-r rows of G_A. For those A_{r-1} which generate the same linear system we can choose the same added row. Thus we obtain only one row for each group of such A_{r-1}. If we now form

the matrix H_A of all these rows and write these under G_A, we obtain a solution matrix $B_A = \binom{G_A}{H_A}$.

The matrix B_A contains no superfluous rows, for among its rows we have only those of G_A, which are orthogonal to A and are therefore essential, and likewise only necessary rows have been added to get H_A. B_A is therefore a solution basis, and each solution basis must be obtainable in this way.

H_A is called the <u>extreme part</u>, G_A the <u>central part</u> of the basis; the latter is suppressed for r=n. The system $xA \in \underline{V}$ is completely solved by $x \in \underline{W}\binom{G}{H}$. All our previous results are contained in this proof; for example, the L_A given for the special A at the beginning of §6 can be deduced in this way.

46. <u>Number of rows</u>. Since each A_{r-1} corresponds to a combination of r–1 columns of A, there are at most $\binom{m}{r-1}$ A_{r-1}, and clearly this number will be obtained if and only if each r–1 columns of A are linearly independent. Thus B has at most $n-r+\binom{m}{r-1}$ rows, and in order that B have exactly this many rows it is necessary and sufficient that each r columns of A be linearly independent.

On the other hand, by Theorem C3 each basis has at least n rows, since G_{A_0} has n rows. For r=1 or r=0 the basis has exactly n rows. It is conjectured that the number of rows of the basis must be at least n–r+p, where p is the largest number of pairwise non-proportional columns of A. Certainly $p \geq r$. This conjecture amounts geometrically to the statement that p points which do

not lie on the same hyperplane, together generate at least p hyperplanes.* Although this conjecture seems to be true, I haven't been able to prove it even for $r=3$ (that is, in the plane; "hyperplane" is a straight line).

47. Relation between bases. If we wish to form a basis $B_1 = \binom{G_1}{H_1}$ of A by taking linear combinations of the rows of a basis $B = \binom{G}{H}$ of A, then we know by 17 that G_1 can be formed from G through left multiplication by a non-singular matrix M_G. Furthermore, each row of H_1 can be represented as a linear combination of the rows of G and the corresponding (i.e., belonging to the same A_{r-1}) row of H, in which the latter has a non-zero coefficient. In other words: each H-row is determined modulo G except for a factor.

Thus we can represent the rows of B_1 in the form $M_B B$ where

$$M_B = \begin{bmatrix} M_G & 0 \\ M_{GH} & M_H \end{bmatrix}, \quad |M_G| \cdot |M_H| \neq 0,$$

and M_H is a diagonal matrix. In other words, we have the relation

$$\text{V} \qquad \binom{G_1}{H_1} = \begin{pmatrix} M_G & 0 \\ M_{GH} & M_H \end{pmatrix} \begin{pmatrix} G \\ H \end{pmatrix}.$$

A matrix M_B of this type is determined by B_1 and B, hence for a given B, B_1 and M_B correspond to each other uniquely.

Conversely, if the product $\overline{M}_B B$ is a basis of A, then $\overline{M}_B B$ must be equal to some $M_B B$, so $(\overline{M}_B - M_B)B = 0$. Thus $\overline{M}_B = M_B + \overline{B}$, where

---

* See Motzkin, The lines and planes connecting points of a finite set, Trans. Am. Math. Soc., May 1951, for a proof of this conjecture.

$\bar{B}$ is a square matrix such that $\bar{B}B = 0$. Hence $\bar{B} = \bar{B}_1\bar{B}_2$, where $\bar{B}_1$ is arbitrary and $\bar{B}_2$ is a G of B. In other words, a necessary and sufficient condition that $\bar{M}_B B$ be a basis of A is that $\bar{M}_B = M_B + \bar{B}_1\bar{B}_2$, where M_B has the form described above and $\bar{B}_2$ is a G of B.

48. <u>Special V</u>. For each $\underline{V}$, $\underline{W}_V\binom{G}{H}$ yields all solutions of II. However, $\underline{W}_V$ can contain $\underline{E}_0$ as a factor one or more times for the H-rows. The factors of $\underline{\dot{W}}_V$ corresponding to the G-rows are of course $\underline{E}_{-0+}$. Thus all rows of H are necessary when we consider the totality of $\underline{V}$, but may not be for a single one.

Instead of considering all $\underline{V}$, one can consider only those $\underline{V}$ which are orthants (products of $\underline{E}_{-0}$ and $\underline{E}_{0+}$) and observe that all rows of H are necessary for the totality of these $\underline{V}$. For if h is a row of H, then hA belongs to some closed orthant $\underline{V}$, so for this $\underline{V}$ h is a solution and consequently not superfluous.

If in II we replace each relation $xa \leq 0$ by $x \cdot -a \geq 0$ and each equation $xa = 0$ by both inequalities $xa \geq 0$, $x \cdot -a \geq 0$, we get an equivalent system with only $\lesseqgtr$ and $\geq$ signs, which can be written in the form $x(A,-A) \in \underline{\bar{V}}$ where $\underline{\bar{V}}$ is a product of $\underline{E}_{-0+}$ and $\underline{E}_{0+}$. If A has m columns, there are 2^{2m} distinct $\underline{V}$, likewise 2^{2m} distinct $\underline{\bar{V}}$.

49. <u>Orthogonalization</u>. As is well known, the rows of a G can be successively orthogonalized and normalized, so that $GG' = E_{n-r}$, i.e., the rows of G have square sum 1 and are pairwise orthogonal. We can also select for each H-row one which

is orthogonal to G and has square sum 1; this row is uniquely determined except for multiplication by -1, so nothing more can be said about the product of two such H-rows.

In this way one obtains a normalized basis. A matrix M_B which transforms one normalized basis into another must have the form $\begin{pmatrix} M_G & 0 \\ 0 & M_H \end{pmatrix}$, where $M_G M_G' = E_{n-r}$ and M_H is a diagonal matrix with ± 1 along the diagonal. Moreover, for a fixed basis, every other basis corresponds to precisely one such M_B. Also, just as before, we can add to such an M_B an arbitrary $\bar{B}$ with $\bar{B}B = 0$.

§9. Fixed $\underline{V}$.

50. Shortest representation. If only a single $\underline{V}$ is considered, then all rows b in the expression of the general solution $x \in \underline{W}_B B$ which correspond to $\underline{W}_B$ components $\underline{E}_0$ (i.e., where neither $b \in \underline{L}$ nor $-b \in \underline{L}$) can be omitted. We can also replace those b which have coefficient ranges $\underline{E}_{-0}$ by $-b$. This of course changes only the rows of H.

Thus from B we get a matrix $C = C_A = \binom{C_1}{C_2}$, such that

$$\text{VI} \qquad x = uC_1 + vC_2, \quad u \text{ arbitrary}, \quad v \geq 0,$$

and conversely each such expression is a solution of II. We suppose that any superfluous rows have been omitted, so that the number of rows of C cannot be reduced. Then

Theorem E2. The solution domain $\underline{L}$ of a closed homogeneous system is the sum of a linear space (uC_1) and a convex polyhedral cone (vC_2).

51. Investigation of C_1. If both $x = uC_1+vC_2$ and $-x$ are in $\underline{L}$, then $uC_1-x = -vC_2$ is in $\underline{L}$. Now the single summands of vC_2 are in L since $v \geq 0$. By adding all of these summands except one to $-vC_2$, we see that non-positive multiples of the rows of C_2 are in $\underline{L}$, a contradiction unless each component of v is zero.

Hence if both x and $-x$ are solutions, then x has the form uC_1, and obviously expressions of the form uC_1 have this property. But these x are solutions of the equation system $xA^* = 0$, where A^* contains all those columns of A which do not correspond to identities of II. Since all superfluous rows were omitted, C_1 is some G_{A^*}, and its number of rows q is uniquely determined by A and $\underline{V}$.

It is also easy to see that the convex cone vC_2 is line-free.

52. Investigation of C_2. Of the remaining x, those which are not positively proportional to a row of C_2 modulo C_1 can be split into two summands which are not positively proportional to each other modulo C_1. As one of these summands one can take a row of C_2 with its (non-zero) coefficient. If this row were positively proportional to the rest of the expression for x, it would be superfluous.

Moreover a row c of C_2 cannot be represented as a sum

$$(u_1C_1+v_1C_2) + (u_2C_1+v_2C_2) \qquad v_1,v_2 \geq 0$$

of two summands in which other rows of C_2 appear. For otherwise one could deduce an equation of the form $\lambda c = uC_1+vC_2$, $v \geq 0$, where c does not appear in vC_2; for $\lambda > 0$, c would again be

superfluous; for $\lambda < 0$, $-c$ would be in $\underline{L}$ and c would correspond to a sign range $\underline{E}_{-0+}$; finally, for $\lambda = 0$, we could bring another row of C_2 to the left and repeat the argument.

In short we see that the rows c (which are "extreme" solutions of II) have certain characteristic properties and are uniquely determined modulo C_1 except for a positive proportionality factor. It follows that the number of rows of C_2 as well as of C_1 are invariants independent of the way C is produced.

Let $C^{(1)}$ be another such C_A. Again we can ask for the multipliers M_0 with $C^{(1)} = M_0 C$. Since the rows of the part $C_1^{(1)}$ of $C^{(1)}$ corresponding to C_1 can be expressed as combinations of the rows of C_1, M_0 has the form

$$\begin{bmatrix} M_1 & 0 \\ M_{12} & M_2 \end{bmatrix}, \qquad |M_1|\,|M_2| \neq 0$$

where M_2, by the remarks above, is a diagonal matrix with positive elements on the diagonal.

53. Positive representations. For a single $\underline{V}$ we can find a matrix D such that $\underline{V}_{0+}D = \underline{L}$, where $\underline{V}_{0+}$ is the principal orthant; in other words, each solution can be written as a non-negative linear combination of the rows of D.

If we have such a D, clearly we can add to it as new rows arbitrary solutions of II. Thus we will be interested in only those D which contain no superfluous rows.

In general there will be some rows of D which remain solutions after multiplication by -1 and others for which this is not the

case; accordingly we write $D = \binom{D_1}{D_2}$. If $x = vD$, $v \geq 0$, has coefficient range $\underline{E}_{-0+}$, then by adding to $-vD$ all summands of vD except one we see that only rows of D_1 appear in vD. The set of these x, that is, the set of all uC_1 of the preceding paragraph, is therefore $V_{0+}D_1$.

The extreme x, which are not decomposable modulo C_1, must be positively proportional to certain rows of D_2 modulo C_1. Since these extreme x and C_1 (or D_1) suffice for the representation of all x, and D_2 should contain no superfluous rows, D_2 is any C_2, and we have only to find D_1.

54. Number of rows. If the number of rows—and hence the rank—of C_1 is q, i.e., uC_1 is a linear space of dimension q, then D_1 must have at least q linearly independent rows. But since the decomposition of each uC_1 into these q rows is uniquely determined, and hence for certain uC_1 negative coefficients will appear, D_1 cannot have only q rows. We will not be concerned with the investigation of all D_1, but will show only that there is a D_1 with $q+1$ rows.[1]

55. Construction of D_1. If there is such a D_1 of $q+1$ rows $d_1, \cdots, d_{q+1}$ where $d_1, \cdots, d_q$ are linearly independent, let $d_{q+1} = \sum_1^q \lambda_i d_i$. Each point of our linear space has the form

$$\sum_1^{q+1} \mu_i d_i, \quad \mu_i \geq 0;$$

[1] The other D_1 can be obtained as follows: one splits the linear space uC_1 into finitely many complementary linear subspaces; each such decomposition corresponds to a decomposition of q into a sum of integers. Then one determines for each subspace a D_1 in the way given in the text, and writes these together as a matrix.

from this we see that if some $\lambda_k > 0$, we could have replaced d_{q+1} by $d'_{q+1} = \sum_{i \neq k} \lambda_i d_i$, and written

$$\sum_{i=1}^{q+1} \mu_i d_i = \sum_{i \neq k, q+1} \mu_i d_i + \mu_{q+1} d'_{q+1} + (\mu_k + \lambda_k \mu_{q+1}) d_k.$$

By repeating this process we reach the conclusion that all $\lambda_i \leq 0$. If some $\lambda_i = 0$, we could not represent $-d_i$. Thus all $\lambda_i < 0$. It follows that $\sum_{1}^{q+1} \chi_i d_i = 0$. $\chi_i > 0$, and each q rows of D_1 are linearly independent.

Conversely we can form from each C_1 arbitrarily many D_1: simply add the row $\sum_1^q \lambda_i d_i$ where the λ_i are negative. If we write $Q_\ell = \begin{pmatrix} E_q \\ \ell \end{pmatrix}$ where ℓ is the row of λ_i, then $D_1 = Q_\ell C_1$, and for each other $D_1^{(1)}$ we have $D_1^{(1)} = Q_\ell^{(1)} C_1^{(1)}$ for some $C_1^{(1)}$. Moreover each $C_1^{(1)}$ has the form $M_o C_1$, with $|M_o| \neq 0$.

Accordingly, each D_1 of q+1 rows can be obtained from a given C_1 through left multiplication by a matrix of the form

$$Q = \begin{pmatrix} E_q \\ \ell \end{pmatrix} M_o, \quad |M_o| \neq 0, \quad \ell < 0.$$

These are all those matrices of q columns and q+1 rows such that a single linear dependence with positive coefficients exists among the rows. Such matrices will be referred to as <u>vertical simplex matrices</u> (the transposes are <u>horizontal simplex matrices</u>).

The matrix

$$\begin{bmatrix} 1 & 0 & . & . & . & 0 & 0 \\ -1 & 1 & . & . & . & 0 & 0 \\ 0 & -1 & . & . & . & 0 & 0 \\ & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & . & . & . & -1 & 1 \\ 0 & 0 & . & . & . & 0 & -1 \end{bmatrix}$$

is a vertical simplex matrix for which all $\chi_1 = 1$; the same is true of the matrix $\begin{pmatrix} E_q \\ -1 \end{pmatrix}$.

56. Supplementary remarks. Geometrically (cf. §14) Q represents a simplex containing the origin in its interior. The μ_1 can be regarded as a kind of simplex coordinate, which for $\mu_1 = 1$ are related to those known by this name as an arithmetic proportion is to a geometric one: they are determined up to an additive constant, not a factor. (The constant must be chosen so that the μ_1 remain non-negative). One can also speak of a projection of the principal orthant of the space $\underline{R}_{q+1}$ onto a suitable supporting hyperplane through the origin.

We look again for the multipliers for our D which let $D^{(1)}$ be another such D_A. In each case we can find a multiplier M_D with $D^{(1)} = M_D D$, which has the form $\begin{pmatrix} M_1 & 0 \\ M_{12} & M_2 \end{pmatrix}$, $|M_1||M_2| \neq 0$, where M_2 contains positive elements on the diagonal, zeros elsewhere, and the rows of M_1 correspond to those of D_1.

M_1 must effect a transition from one simplex to another which also contains the origin, that is, M_1Q must satisfy the conditions

given above for Q. Let the coefficients in the linear relation between the rows of Q be ν_1 and correspondingly $\nu_1^{(1)}$ for M_1Q. Denote the row of ν_1 by k, that of the $\nu_1^{(1)}$ by $k^{(1)}$. Then $k^{(1)}M_1Q = 0$. Now $k^{(1)}M_1$ is a row $\bar{k}$ and since there is only one linear relation between the rows of Q, we must have $\bar{k} = \mu k$, i.e., $k^{(1)}M = \mu k$, which also clearly suffices. Thus the necessary and sufficient condition for M_1 is that $(k^{(1)},\mu)\begin{pmatrix}M_1\\-k\end{pmatrix} = 0$ for a suitable $k^{(1)} > 0$. In other words, $\begin{pmatrix}M_1\\k\end{pmatrix}$ or $\begin{pmatrix}M_1\\-k\end{pmatrix}$ must be a simplex matrix.

Also it is easy to see that the elements of M_1 and M_{12} can all be taken as simplex coordinates ≥ 0.

§10. Inhomogeneous systems.

57. Homogenizing. We proceed to the consideration of inhomogeneous systems Σ: $xA+c \in \underline{V}$, still excluding strict inequalities. We shall handle such systems by adding to the n unknowns $\xi_1,\cdots,\xi_n$ (the coordinates of x) another unknown ξ, regarding the constant term as the coefficient of ξ, where of course the inhomogeneous relation $\xi = 1$ must be added.

The relation $\xi = 1$ represents a hyperplane in $\underline{R}_{n+1}$, not passing through the origin, on which the solutions must lie. If we omit the relation $\xi = 1$, this corresponds geometrically to joining the solution points with the origin, which plays here fundamentally (in a projective sense) only the role of some point outside the original space $\underline{R}_n$.

If we project the solution points from a new origin not by lines but by closed half-lines, then we have replaced the condition $\xi = 1$ by $\xi \in \underline{R}_{0+}$, wherewith the system assumes the form of those already studied.

We can write the general solution of this new system Σ' in the form $x = wL$, $w \in \underline{W}$. Since we will restrict ourselves now to the consideration of a given $\underline{V}$, we can omit those rows of L which take zero coefficients and multiply those rows whose coefficient range is $\underline{E}_{-0}$ by -1, so that $\underline{W}$ is a product of $\underline{E}_{0+}$ and $\underline{E}_{-0+}$ factors.

58. The three constituents. We divide L into three submatrices L_1, L_2, L_3, so that the general solution has the form $x = w_1 L_1 + w_2 L_2 + w_3 L_3$.

L_1 consists of those rows having coefficient ranges $\underline{E}_{-0+}$; each of these rows must have its last component zero, since they and their negatives satisfy $\xi \in \underline{E}_{0+}$. Then w_1 is arbitrary.

We include in L_2 those rows which correspond to $\underline{E}_{0+}$ and end in zero. Thus $w_2 \geq 0$. The rows of L_1 and L_2 are not solutions of the original system in which $\xi = 1$.

Finally we include in L_3 those rows which do not end in zero, whose last coordinates are therefore positive (since the rows themselves are solutions) and can be made 1 by multiplication by a suitable positive factor. Then $w_3 \geq 0$. If we now consider the relation $\xi = 1$, we see also that the sum of the coordinates of w_3, which we write as $w_3 \cdot 1$, should be 1. Otherwise w_3 is arbitrary. Thus

Theorem 05. The general solution of an inhomogeneous relation system I has the form

VII $\qquad x = w_1 L_1 + w_2 L_2 + w_3 L_3, \quad w_2, w_3 \geq 0, \quad w_3 \cdot 1 = 1.$

We think of the $(n+1)^{st}$ coordinate as suppressed, so that $x \in \underline{R}_n$. Since the L for Σ' contained no superfluous rows, all rows of L are necessary for Σ.

For each row ℓ of L_1, and each relation $xa+\gamma \in \underline{E}$ of Σ, the product $\omega(\ell,o)\begin{pmatrix} a \\ \gamma \end{pmatrix} = \omega \ell a$ is in $\underline{E}$ for each ω; if $\underline{E} \neq \underline{E}_{-0+}$, then $\ell a = 0$.

59. Polyhedral sets. The constituent w_1L_1 represents the points of the linear space generated by the rows of L_1; w_2L_2 is (cf. §51) a line-free convex polyhedral cone; finally w_3L_3 is the convex polyhedron generated by the rows of L_3. Since $\underline{L}$ is an arbitrary closed polyhedral set, we see:

Theorem E3. *Every closed polyhedral set can be represented as the sum of a linear space, a convex polyhedral cone, and a convex polyhedron.*[1]

One or two of the three sets in this decomposition may not appear.

If $x+\lambda y$ is a solution of Σ for each $\lambda \geq 0$, then $(x+\lambda y, 1)$ is a solution of Σ', and also $\lim\limits_{\lambda^{-1}\to 0} (\frac{x}{\lambda}+y, \frac{1}{\lambda}) = (y, 0)$. Thus y has the form $w_1L_1+w_2L_2$, $w_2 \geq 0$. Consequently $x+w_1L_1+w_2L_2$, $w_2 \geq 0$, represents the biggest convex cone with vertex x which is contained in $\underline{L}$.

If $\underline{L}$ itself is a convex cone, then each point of the polyhedron w_3L_3 can be chosen as vertex, and therefore w_3L_3 must reduce to a point, since no rows of L_3 are superfluous; that is, L_3 has only one row. (We must assume at least one row in L_3, even though

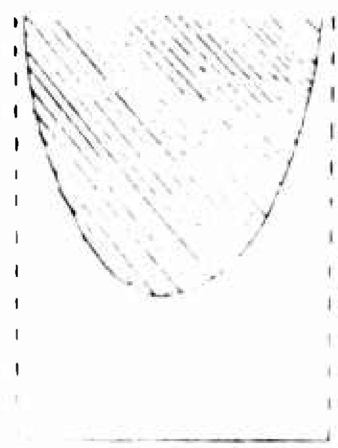
[1] In general, each wholefaced convex set is representable as a sum of a linear space and a line-free convex set, and this (up to parallel displacement) in a unique way; the decomposition above into a bounded convex set and a convex cone cannot be accomplished in general.

it is the zero row, in order to satisfy the condition $w_3 \cdot 1 = 1$).
Hence the general solution will be

$$x = w_1 L_1 + w_2 L_2 + L_3, \qquad w_2 \geq 0.$$

It is also easy to see that any vertex of $\underline{L}$ has the form
$x = v_1 L_1 + L_2$.

60. Geometric remarks. The projection of a convex set onto a line is evidently convex. However, the projection of a closed convex set is not in general closed (see the figure). But for a polyhedral set this too is valid. For if we form, for a fixed vector a, the product $xa = w_1 L_1 a + w_2 L_2 a + w_3 L_3 a$, then we get an expression of the form $\Sigma w_i \ell_i a$, where the $\ell_i a$ are fixed numbers and each w_i is arbitrary except that certain w_i are positive and some of them have sum 1; obviously this expression assumes its extreme values.



From the discussion of the last paragraph we perceive the geometric function of the solution matrix L of a homogeneous system. For this purpose we interpret the variables as homogeneous coordinates in a projective space $\underline{R}_{n-1}$; this is decomposed by the finitely many hyperplanes corresponding to the relations P into finitely many polyhedral spatial parts, and the solution domains of Σ must be formed from such. The finitely many "corners" (independent boundary manifolds) namely the rows of H_A, combined with the linear manifold generated by G_A, suffice for the characterization and representation of all these polyhedral parts and

solution domains. Incidentally, G and H correspond to the first and second cases in the proof of Theorem C3.

61. <u>Strict</u> <u>inequalities</u>. The homogeneous relation system $xA \in \underline{V}$ is solved by Theorem C2 in the form $x = wA^{(1)}$, $w \in \underline{W}$. Let a be a column of A which corresponds to an inequality. If we write $x = wA^{(1)} = \Sigma\omega a^{(1)}$, where the $a^{(1)}$ are rows of $A^{(1)}$ and ω ranges over the coordinates of w, then we have $xa = \Sigma\omega a^{(1)}a$; moreover, each $\omega a^{(1)}a$ which is different from zero has the sign demanded by the inequality. In order that $xa = 0$, we must have each $\omega a^{(1)}a = 0$; i.e., for $a^{(1)}a \neq 0$, ω must vanish.

The replacement of an inequality by the corresponding equation means that certain coefficients of x must vanish. If on the other hand we replace an inequality by the corresponding strict inequality, then at least one of these ω-coordinates must be different from zero. Thus we have extended the parametric representation of the general solution to those finite systems of homogeneous relations involving the signs $=$, $\geq$, $\leq$, $>$, $<$.

Suppose now we are faced with the most general case of an inhomogeneous system in which strict inequalities are admis[illegible]. We can homogenize the system by adding a new unknown and han[illegible] it as above, or we can replace all >0 and <0 by $\geq\eta$ and $\leq-\eta$, add the relation $\eta > 0$, and then homogenize it. This reduces the number of strict inequalities to one ($\eta > 0$).

Geometrically, replacing some of the inequalities by strict inequalities excludes at most certain pieces of the boundary of the convex set of all solutions; in this way the most general polyhedral set is obtained.

CHAPTER III. SOLVABILITY CRITERIA

§11. Minimal unsolvable systems.

62. Definition. In §12 we will obtain conditions for the unsolvability of general relation systems.

An unsolvable system remains unsolvable after the addition of other relations. Consequently the unsolvability of a system depends on the unsolvability of a subsystem, perhaps this in turn on another, etc.; eventually we arrive at a minimal unsolvable system (m.u.s.), i.e., one whose proper subsystems are solvable.

Each unsolvable system contains at least one m.u.s., and consequently we restrict ourselves first to these. Our principal goal is to prove Theorem D1. For the time being, we note that a m.u.s. can consist of a single relation, for example $1 \leq 0$, and never contains identities.

63. $\underline{L}$ and $\underline{\bar{L}}$. Suppose $xA+c \in \underline{V}$ is a m.u.s., where A has n rows and m columns. After the removal of any relation P: $xa+\delta \in \underline{E}$, we are left with a solvable system $\Sigma - P$, having general solution y; if we replace each relation by the corresponding closed relation, the system remains solvable, and its general solution $\bar{y}$ (Theorem C5) has the form $\bar{y} = w_1 L_1 + w_2 L_2 + w_3 L_3$, $w_2, w_3 \geq 0$, $w_3 \cdot 1 = 1$. Let $\underline{L}$ be the set of all y, $\underline{\bar{L}}$ the set of all $\bar{y}$. Since $\underline{L}$ is wholefaced, $\underline{\bar{L}}$ is the closure of $\underline{L}$ (Theorem F2).

Now $ya+\delta$ assumes no value in $\underline{E}$, as otherwise the system would be solvable. Hence if $\bar{y}a+\delta$ for $\bar{y} = y_0$ assumes a value $\alpha \neq 0$ of E, then, since y can be taken arbitrarily close to y_0, $ya+\delta$ could be brought arbitrarily close to α, which is incompatible with the

unsolvability of Σ. We see therefore:

$\bar{y}a+\gamma$ can assume at most the value 0 from the sign range $\underline{E}$.

Analogously one sees that $\bar{y}a+\gamma$ cannot take on both positive and negative values.

64. The z-corner. We have observed (60) that $\underline{L}a$, the projection of $\underline{L}$ onto a, is a closed convex subset of $\underline{R}_1$; hence $\underline{L}a+\gamma$ is also. It follows that there is precisely one number ζ for which

(1) $\quad |\bar{y}a+\gamma| \geq |\zeta|$,

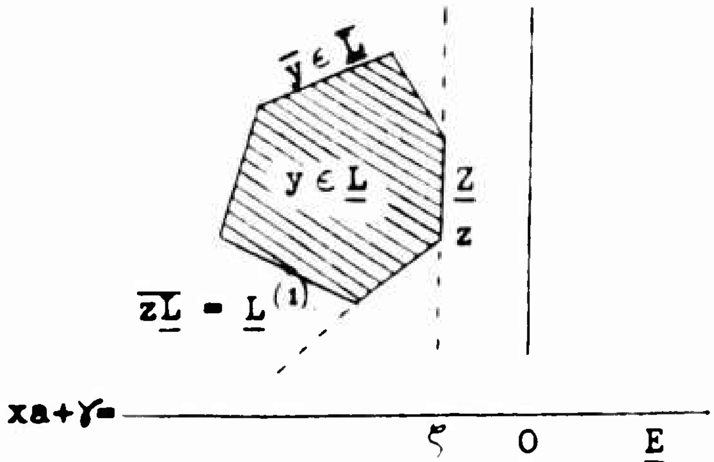
and, if $\zeta \neq 0$,

(2) $\quad sg(\bar{y}a+\gamma) = sg\,\zeta$.

Let $\underline{Z}$ be the (non-empty) set of all $\bar{y}$ for which the extreme value ζ is reached, and let z be a point of $\underline{Z}$, so that

(3) $\quad za+\gamma = \zeta$.

(See the figure, although it represents a case which is impossible for a m.u.s., as we shall see.)



Now consider the z-corner of Σ-P, that is, the system of all relations $P^{(1)}$ of Σ-P for which $P^{(1)}(z) = 0$; since $z \in \underline{L}$, by Theorem E1 the convex cone $\overline{z\underline{L}}$ is the same as the solution domain $\underline{L}^{(1)} = \underline{L}(\Sigma^{(1)})$.

65. $\underline{L}$ as a convex cone. Can the system $(\Sigma^{(1)}, P)$ be solvable? That is, is there a point $p \in \underline{L}^{(1)}$ such that $pa+\gamma \in \underline{E}$? As a point of $\underline{L}^{(1)}$, p has the form $z+\lambda(y-z)$, $\lambda > 0$, or for $z \in \underline{L}$,

$p = z$; however, the latter is excluded, as otherwise z would be a solution of Σ. It follows from $p = z+\lambda(y-z)$ that $pa+\gamma = (1-\lambda)\zeta+\lambda(ya+\gamma)$.

If $\zeta \neq 0$, then by (1) and (2), $|pa+\gamma| \geq |\zeta|$ and $sg(pa+\gamma) = sg\,\zeta$. Hence $pa+\gamma$ cannot lie in $\underline{E}$ in this case.

If $\zeta = 0$, then $pa+\gamma = \lambda(ya+\gamma)$, $\lambda > 0$, so if $pa+\gamma \in \underline{E}$, also $ya+\gamma \in \underline{E}$, and Σ would be solvable.

The question asked above is answered in the negative: the system $(\Sigma^{(1)}, P)$ is unsolvable and is therefore identical with Σ. We see that z reduces all the linear forms other than the excluded one to zero.

From $\Sigma^{(1)} = \Sigma - P$ it follows that $\overline{z\underline{L}} = \underline{L}^{(1)} = \underline{L}$, so we have proved the theorem:

<u>Theorem E4. The solution domain L of a system arising from a</u> m.u.s. <u>by the omission of a single relation is a convex cone.</u>

66. <u>$\underline{Z}$ as a linear manifold.</u> Since $\underline{\overline{L}}$ results from $\underline{L}$ by adding hyperplanes passing through z, $\underline{\overline{L}}$ is also a convex cone, and by 59 the polyhedral part L_3 has only one row. Then

$$\overline{y} = w_1L_1+w_2L_2+L_3, \quad w_2 \geq 0$$

and $\overline{y}a = w_1L_1a+w_2L_2a+L_3a$. By 63, $\overline{y}a+\gamma$, and hence $\overline{y}a$, does not range over all values, so neither does w_1L_1a. But w_1 is arbitrary, so L_1a must be zero. By 58, since Σ contains no identities, each row of L_1 multiplied by each column of A gives zero, $L_1A = 0$.

Conversely each row ℓ for which $\ell A = 0$ does not change the value of $\overline{y}A$ when it is added to $\overline{y}$, whence, as one deduces from properties of the summands of $\overline{y}$, ℓ must be a linear combination of

rows of L_1. Since L_1 has no superfluous rows, L_1 is some G_A and has $n-r$ rows, where r is the rank of A. For $r = n$, $L_1 = 0$. It can also be seen directly that L_1 is some G_{A^1}, whence the rank $r^{(1)}$ of $A^{(1)}$ is equal to r.

By 59, each element $z \in \underline{Z}$ has, as a center of $\underline{L}$, the form $z = w_1L_1+L_3$, and since, by 64,(3),

$$(w_1L_1+L_3)a = L_3a = \zeta - \gamma$$

is not dependent on w_1, $w_1L_1+L_3$ always represents a z. Thus, since L_1 has $n-r$ linearly independent rows and w_1 is arbitrary, $\underline{Z}$ is a linear manifold of dimension $n-r$, which satisfies $m-1$ of the m equations corresponding to the relations of A; and indeed precisely $m-1$, unless $\zeta = 0 \in \underline{E}$.

67. <u>Strict inequalities</u>. In the exceptional case $\zeta = 0$, $xA+c \in \underline{V}$ will itself be solvable if each relation is replaced by its corresponding closed relation.

Then there must be at least one strict inequality $xa+\gamma \in \underline{V}$. This inequality represents an open half space; if this half space is replaced by a hyperplane contained in it, i.e., the inequality is replaced by an equation $xa+\delta = 0$, the system remains a m.u.s. In fact the system remains unsolvable since only a sharpening has occurred. But the solvable system obtained by omitting the inequality was a z-cone, where z did not belong to the open half space; if such a cone intersects the half space, it must also intersect the hyperplanes parallel to the boundary.

If we proceed in this way with each strict inequality, we finally get a closed m.u.s. with the same matrix A.

68. *The rank of A is m–1.* Consequently the exceptional case can be disregarded, and we conclude as follows:

The omission of each relation in sequence yields m linear manifolds which are independent of each other, since each m–1 satisfy an equation which the m^{th} does not. By selecting n–r independent points from one of the manifolds, and one each from the remaining, we see that the space generated by the manifolds taken together has dimension at least m–1+n–r. Since the total space is of dimension n, then $n \geq m-1+n-r$, i.e., $r \geq m-1$.

But since m independent linear forms have a solution, $r < m$ and therefore $r = m-1$. Also, since $r^{(i)} = r$, the columns of each of the m subsystems $A^{(i)}$ are linearly independent.

Theorem D1. *The rank of the coefficient matrix of a m.u.s. is one less than the number of relations in the system.*

The matrix $\binom{A}{c}$ has rank m–1 or m and thus the m.u.s. are of two kinds, which one can classify further with respect to $\underline{V}$; the rank m–1 corresponds to the exceptional case $\xi = 0$.

§12. *Unsolvability.*

69. *The combination principle.* Let $xA+c \in \underline{V}$ be any unsolvable system; then it contains a m.u.s. $xA^{(i)}+c^{(i)} \in \underline{V}^{(i)}$. We have proved that each i of the i+1 columns of $A^{(i)}$ are linearly independent. For $i \leq k \leq r$, we can add to such a set of i columns k–i columns of A not in $A^{(i)}$ such that the k columns are linearly

independent. In this way we obtain from $A^{(1)}$ a matrix $A^{(2)}$ with k+1 columns and a corresponding system $xA^{(2)}+c^{(2)} \in \underline{V}^{(2)}$. We write this as

$$xA^{(2)} = v^{(2)}-c^{(2)}; \quad v^{(2)} \in \underline{V}^{(2)}.$$

Thus for each $v^{(2)} \in \underline{V}^{(2)}$ we have an unsolvable system of equations, whose matrix $\begin{pmatrix} A^{(2)} \\ v^{(2)}-c^{(2)} \end{pmatrix}$ must therefore have rank k+1.

Let any k linearly independent rows of $A^{(2)}$ form a matrix $A^{(3)}$. Such $A^{(3)}$ exist since $A^{(2)}$ has rank k, and the corresponding sub-determinants of order k of distinct $A^{(3)}$ are proportional to each other (Theorem A4). Then also $\begin{pmatrix} A^{(3)} \\ v^{(2)}-c^{(2)} \end{pmatrix}$ has rank k+1, i.e. $\begin{vmatrix} A^{(3)} \\ v^{(2)}-c^{(2)} \end{vmatrix} \neq 0$. Thus the equation $\begin{vmatrix} A^{(3)} \\ v^{(2)} \end{vmatrix} = \begin{vmatrix} A^{(3)} \\ c^{(2)} \end{vmatrix}$, $v^{(2)} \in \underline{V}^{(2)}$, is unsolvable. This equation is a linear combination of the original equations, whose coefficients are certain subdeterminants of A. By 25, the linear form $\begin{vmatrix} A^{(3)} \\ v^{(2)} \end{vmatrix}$ for $v^{(2)} \in \underline{V}^{(2)}$ ranges over some one dimensional sign range $\underline{E}$.

Thus a necessary and sufficient condition that an inhomogeneous system of rank $\geq$ k be unsolvable is that at least one of the $\begin{vmatrix} A^{(3)} \\ c^{(2)} \end{vmatrix}$ formed in this way not lie in the associated $\underline{E}$. We formulate this in the combination principle:

Theorem D2. <u>A contradiction free of unknowns can be derived from every unsolvable system by a linear combination of relations of the system.</u>

Such a contradiction is of the form $\alpha < \alpha$, or one of $\alpha \leq \beta$, $\alpha = \beta$, $\alpha < \beta$ where really $\alpha > \beta$.

70. Definite dependence. A linear relation between the columns of a matrix in which the coefficients have the same sign throughout, but are not all zero, will be called a definite dependence of the columns (or the matrix) (concerning "definite" and "strictly definite", see 15).

From the definition of a simplex matrix (55), a matrix of i rows and i+1 columns is a simplex matrix if there is precisely one strict linear dependence between its columns; the coefficients of this dependence are then proportional to the i^{th} order subdeterminants of A multiplied alternately by +1 and -1. Thus we see

Theorem B1. A matrix of i rows and i+1 columns is a simplex matrix if and only if its i+1 i^{th} order subdeterminants alternate in sign.

Suppose we are given a m.u.s. of the form $xA \in \underline{V}_+$ of r+1 strict homogeneous inequalities, which can also be written as $xA > 0$. By Theorem D1, A has rank r and so there is just one linear relation (except for a factor) between the r+1 columns of A; moreover, since each r columns of A are linearly independent, none of the coefficients χ_i of this relation is zero.

On the other hand, by Theorem D2, a linear combination of the inequalities produces a contradiction free of unknowns; the coefficients of this linear combination must be the χ_i of the linear dependence, and thus are proportional to r^{th} order subdeterminants (from some r rows of A) alternately multiplied by +1 and -1.

Now $\Sigma \chi_i \underline{E}_+$ is $\underline{E}_0$ if all $\chi_i = 0$, $\underline{E}_+$ (resp. $\underline{E}_-$) if all $\chi_i \geq 0$ (resp. ≤ 0) but not all $\chi_i = 0$, and $\underline{E}_{-0+}$ in all remaining cases. Since $0 \in \Sigma \chi_i \underline{E}_+$ should be a contradiction and since no χ_i is zero, all χ_i must be either positive or negative. Hence the subdeterminants alternate in sign.

71. Principal criterion. If now we consider, instead of a m.u.s., any unsolvable system $xA > 0$, then we have

Theorem D3. *A system of the form $xA > 0$ is solvable if and only if no $i+1$ linearly dependent columns of A contains a simplex matrix of i rows and $i+1$ columns.*

For $i = 0$ this obviously should mean that no column of A contains only zeros.

This principle can also be stated as follows:

Theorem D4. *$xA > 0$ is solvable if and only if for each horizontal simplex matrix B of A there is a row s for which the determinant $\begin{vmatrix} B \\ s \end{vmatrix} \neq 0$.*

For if $\begin{vmatrix} B \\ s \end{vmatrix} = 0$ for all s, we would have the case forbidden by Theorem D3.

To apply this criterion to a matrix A, we need to know only the signs of its subdeterminants (including its elements). We will assign to a simplex matrix of i rows and $i+1$ columns that sign which the subdeterminant of its first i columns has.

72. Corollary. We now investigate the special case where the coefficient matrix has the form (E,A), i.e., the system is $xA > 0$, $x > 0$.

Each submatrix of (E,A), after possibly a permutation of rows, has the form $B^* = \begin{pmatrix} E_p & C \\ 0 & B \end{pmatrix}$, or a "subform" of this, e.g., $\begin{pmatrix} C \\ B \end{pmatrix}$; the latter cases are subsumed by the general one with only trivial modifications.

When is $B^* = \begin{pmatrix} E_p & C \\ 0 & B \end{pmatrix}$ a simplex matrix? Let the number of rows of B^* be l; then B has $l-p$ rows and $l-p+1$ columns. If we omit one of these $l-p+1$ columns, the remaining determinant of B^* has the value of a subdeterminant of highest order of B: thus B itself must be a simplex matrix. Omission of the last column shows that $\mathrm{sg}B = \mathrm{sg}B^*$.

But if we omit one of the first p columns, say the k^{th}, and if t_k is the k^{th} row of C, then the remaining determinant is $(-1)^{p-k}\begin{vmatrix} t_k \\ B \end{vmatrix} = (-1)^{l-k}\begin{vmatrix} B \\ t_k \end{vmatrix}$. The signs alternate with k, hence in each case the sign of $\begin{vmatrix} B \\ t_k \end{vmatrix}$ is independent of k. Since omission of the p^{th} column gives a determinant $\begin{vmatrix} t_p \\ B \end{vmatrix}$ of sign $(-1)^{l-p+1}$ $\mathrm{sg}B^*$, then

$$\mathrm{sg}\begin{vmatrix} B \\ t_p \end{vmatrix} = (-1)^{l-p}\mathrm{sg}\begin{vmatrix} t_p \\ B \end{vmatrix} = -\mathrm{sg}B,$$

and all of these signs are opposite that of B. Evidently this condition is also sufficient.

Now let B be a horizontal simplex matrix of A, such that there is no row s of A not in B for which $\mathrm{sg}\begin{vmatrix} B \\ s \end{vmatrix} = \mathrm{sg}B$. Then we pick out all s with $\mathrm{sg}\begin{vmatrix} B \\ s \end{vmatrix} = -\mathrm{sg}B$, write them as a p-rowed matrix c over B and add the columns $\begin{pmatrix} E_p \\ 0 \end{pmatrix}$ to get a matrix B^* with l rows and $l+1$ columns. By what we just saw, B^* is a simplex matrix, and

for each other s of A, $\begin{vmatrix} B^* \\ s \end{vmatrix} = 0$, whence the system is unsolvable.

Theorem D5. The system $xA > 0$ is solvable by positive unknowns x if and only if for each horizontal simplex matrix B of A there is a row s with $\operatorname{sg}\begin{vmatrix} B \\ s \end{vmatrix} = \operatorname{sg} B$.

§13. The transposition theorem.

73. Its derivation. We have seen (Theorem D2) that from each unsolvable system a contradiction free of unknowns can be produced by multiplying the relations by certain factors (namely subdeterminants), and adding.

Now each homogeneous system can be written in the form (perhaps after certain relations are multiplied by -1)

VIII $\quad\Sigma$: $xA > 0$, $xB \geq 0$, $xC = 0$,

and we suppose in the sequel that A actually appears. This can be rewritten as

$$x(A,B,C) \in \underline{V} = \underline{V}_+ \cdot \underline{V}_{0+} \cdot \underline{V}_0,$$

where the dimensions of the three factors of $\underline{V}$ are equal to the number of columns of A,B,C respectively.

Since the relations are homogeneous, the contradiction arising through combination will be of the form

$$Ay_1+By_2+Cy_3 = 0, \qquad 0 \notin \Sigma\eta_1\underline{E}_+ + \Sigma\eta_2\underline{E}_{0+} + \Sigma\eta_3\underline{E}_0,$$

where, for example, η_1 runs through the coordinates of y_1. Hence $\Sigma\eta_1\underline{E}_+ + \Sigma\eta_2\underline{E}_{0+} + \Sigma\eta_3\underline{E}_0$ is either $\underline{E}_+$ or $\underline{E}_-$, which can happen only if at least one $\eta_1 \neq 0$ and all η_1 and η_2 different from zero have the same sign, which we can assume positive.

In other words: if VIII is unsolvable, the system

$$\Sigma': \quad (A,B,C)y = 0, \quad y_1 \geq 0, \; y_2 \geq 0, \quad y_1 \neq 0$$

is solvable; the converse is also valid: if VIII is solvable, Σ' is unsolvable, since both together give a contradiction. We shall call Σ' the <u>transposed system</u> since its coefficient matrix is the transpose of that of Σ. We thus have the transposition theorem:

<u>Theorem</u> D6. <u>A system Σ and its transposed system Σ' are complementary in the sense that precisely one of them is solvable</u>.

74. <u>Row-wise interpretation</u>. Geometrically the transposition theorem can be interpreted in many ways. We can consider either the rows or columns of the coefficient matrix as points in cartesian or homogeneous space.

We first think of the rows of (A,B,C) as cartesian coordinates of points. For variable x, x(A,B,C) ranges over all linear combinations of these points, i.e., a linear space $\underline{X}$ through the origin. Those y which satisfy the equations $(A,B,C)y = 0$ of the transposed system form a linear space $\underline{Y}$ which is the orthogonal complement of $\underline{X}$ at the origin. If we choose in particular $B = C = 0$, we see that the two statements "$\underline{X}$ and $\underline{V}_+$ are disjoint" and "$\underline{Y}$ and $\underline{V}_{0+}$ have only the origin in common" are complementary. In other words:

<u>Theorem</u> G1. <u>In each two orthogonally complementary linear spaces through the origin, at least one has points other than the origin in common with the principal orthant, and at most one has points common to the interior of the principal orthant</u>.

This can also be seen directly.

Each $y \neq 0$ in $\underline{V}_{0+}$ can also be interpreted as a coordinate vector of a hyperplane on which the points of X lie, and similarly for $\underline{X}$ and $\underline{Y}$ interchanged. Thus

Theorem G2. Each linear space not intersecting the principal orthant (resp., its interior) lies on a hyperplane not intersecting the principal orthant (resp., its interior).

We note that the principal orthant and the linear space generate a convex set, the hyperplane is some supporting hyperplane through the linear space.

75. Column-wise interpretation in $\underline{R}_n$. If we think of the columns of (A,B,C) as points, then $(A,B,C)y = 0$, $y_1 \geq 0$, $y_2 \geq 0$, $y_1 \neq 0$ for $B = C = 0$ means that the origin lies in the interior or on the boundary of the convex polyhedron $\underline{P}$ generated by A. Or we can say that the columns a of A are definitely dependent.

The statement $xA > 0$ means that x forms an acute angle with each column a of A, i.e., each column and thus all of $\underline{P}$ lies on one side of the hyperplane perpendicular to x at the origin. Since we can take any point p as the origin, we have

Theorem G3. Through each point exterior to a convex polyhedron there is some hyperplane not passing through the polyhedron.

Clearly this is valid for convex sets in general: one need only take the hyperplane perpendicular to $\overline{pq}$ at p, where q is the point of the closure $\underline{\bar{A}}$ of $\underline{A}$ which is closest to p.

The general theorem:

Theorem G4. Through each boundary point of a closed, convex set there is a supporting hyperplane, which can be obtained from the last statement by a limiting process, will be derived geometrically in 78.

76. Column-wise interpretation in R_{n-1}. Finally let the columns of A be homogeneous coordinates in R_{n-1}, with $\xi_n = 0$ interpreted as the hyperplane at infinity. Then we divide the matrix A into two submatrices A_+ and A_-. We put the columns with $\xi_n > 0$ into one group, those with $\xi_n < 0$ into the other (we suppose both appear) and distribute the remaining ones (where $\xi_n = 0$) arbitrarily into both. Accordingly, Σ' assumes the form

$$A_+y_+ + A_-y_- = 0, \; y = \begin{pmatrix} y_+ \\ y_- \end{pmatrix}, \quad y \geq 0, \; y \neq 0.$$

Now $y_+ \neq 0$, since if $y_+ = 0$, also $A_-y_- = 0$ and thus, considering the last coordinate of the product, it would follow that $y_- = 0$; in fact, at least one of the coordinates of y_+ which correspond to columns of A ending in non-zero numbers, is non-zero.

Now A_+y_+ for $y_+ \geq 0$, $y_+ \neq 0$ is some point of the convex set K_+ generated by the columns of A_+, which, thought of as a sum of a convex cone and a polyhedron, is polyhedral. The solvability of Σ' means that K_+ and K_- (defined similarly) are not disjoint, but have namely the point $A_+y_+ = -A_-y_-$ in common.

An x with $xA > 0$ represents the coordinates of a hyperplane which, as one sees easily, separates K_+ and K_-. Thus the transposition theorem gives:

Theorem G5. *Two closed disjoint polyhedral sets can be separated by a hyperplane having a positive distance from both.*

For bounded sets (not necessarily polyhedral, but closed and convex)[1], the hyperplane perpendicular to $\overline{pq}$ at its midpoint, where p and q are two points of K_+ and K_- at minimal distance from each other.

77. *Existence of a separating hyperplane.* In order to prove the general theorem just stated for closed convex sets, which are a positi' distance apart, one needs the statement that two closed convex sets without common interior points can be separated by a hyperplane (neglecting boundary points); then one separates their so-called ε-neighborhoods.

We can deduce this latter statement as follows.

Let the given sets $\bar{A}$ and $\bar{B}$ lie in R_{n-1}, and denote their interiors by A and B; we then think of R_{n-1} as a hyperplane H in R_n. Choose a point p in R_n not in H, and reflect B through p to get a congruent set B^* which lies in a hyperplane H^* parallel to H.

I assert that the set C of all lines joining points of A with poirts of B^* is convex. For suppose z_1 and z_2 are two points of C, and let z_1 be a point on the line from x_1 to y_1, z_2 a point on the line from x_2 to y_2, $x_1, x_2 \in A$, $y_1, y_2 \in B^*$.

[1]The theorem was stated by Minkowski for this case. Only one of the sets needs to be bounded.

Thus

$$z_1 = \lambda_1 x_1 + \mu_1 y_1, \qquad \lambda_1, \mu_1 > 0, \qquad \lambda_1 + \mu_1 = 1$$

$$z_2 = \lambda_2 x_2 + \mu_2 y_2, \qquad \lambda_2, \mu_2 > 0, \qquad \lambda_2 + \mu_2 = 1.$$

Then for arbitrary λ, μ with $\lambda, \mu > 0$, $\lambda + \mu = 1$,

$$\lambda z_1 + \mu z_2 = (\lambda\lambda_1 + \mu\lambda_2)\left(\frac{\lambda\lambda_1}{\lambda\lambda_1 + \mu\lambda_2} x_1 + \frac{\mu\lambda_2}{\lambda\lambda_1 + \mu\lambda_2} x_2\right)$$

$$+ (\lambda\mu_1 + \mu\mu_2)\left(\frac{\lambda\mu_1}{\lambda\mu_1 + \mu\mu_2} y_1 + \frac{\mu\mu_2}{\lambda\mu_1 + \mu\mu_2} y_2\right)$$

$$\lambda\lambda_1 + \mu\lambda_2 + \lambda\mu_1 + \mu\mu_2 = 1$$

Now p does not belong to $\underline{C}$, as otherwise $\underline{A}$ and $\underline{B}$ would have common points. On the other hand, p may belong to the closure $\underline{\bar{C}}$ of $\underline{C}$ and certainly belongs to the boundary of the convex cone $\overline{p\underline{C}}$. If we take (see 78) a supporting hyperplane $\underline{H}_1$ to $\overline{p\underline{C}}$ at p, then $\underline{H}_1$, and consequently its intersection with $\underline{H}$, separates $\underline{\bar{A}}$ and $\underline{\bar{B}}$. The intersection is a hyperplane in $\underline{R}_{n-1}$.

A proof can also be developed along the following lines: for each given set, form one similarly situated and contained in the given set; separate these sets (by 76) by a hyperplane, and let the inner sets converge to $\underline{A}$ and $\underline{B}$; then the hyperplane tends to a limiting position which can be shown to separate $\underline{A}$ and $\underline{B}$.

78. __Existence of a supporting hyperplane__. Theorem G4 is trivial in $\underline{R}_1$ and easily deduced in $\underline{R}_2$ by letting a half line rotate about the given point O. We shall prove the theorem in $\underline{R}_n$, $n > 2$, on the assumption that it is valid for smaller n.

We take a hyperplane $\underline{H}$ through $0 = 0_n$; $\underline{H}$ intersects the given set $\underline{K}$ in $\underline{K}_1$ (also convex). If $\underline{K}_1$ contains 0 in its interior,

then H has the desired property: for it is easily seen that if points of K lay on both sides of H, then 0, contrary to assumption, must be an interior point.

Suppose 0 is on the boundary of K_1: by the induction assumption there is an $(n-2)$-dimensional linear space H_1 in H, so that K_1 lies on one side of H_1. Let G be the line perpendicular to H at 0, G_1 the line in H perpendicular to H_1 at 0. Then we project K onto the plane $E = (G, G_1)$ and get a new convex set K_2.

Since 0 is a boundary point of K_2, there is a supporting line G_2 of K_2 at 0. Since E is perpendicular to H_1, G_2 is perpendicular to H_1. Now G_2 and H_1 generate a hyperplane H_2 which has the desired property; for H_2 projected on E is G_2, so H_2 must be a supporting hyperplane of K.

79. Completely signed matrices. By Theorem G4, the origin is in the interior of a convex set generated by a given set A if and only if points of A lie on both sides of each hyperplane through the origin. Incidentally, Kakeya uses this as a definition of convex generation.

Each hyperplane $xa = \alpha$ which does not pass through the origin ($\alpha \neq 0$) and is parallel to no axis (all coordinates of a are non-zero) does not intersect some unique orthant. Accordingly there is a convex set, which has points in the interior of each orthant except one, such that the origin is exterior to the set.

But if a convex set A has points in the interior of each orthant, then the origin is in the interior of A. In order to see this, consider all points of A with positive first coordinate ξ_1;

in this set there is a point on the positive ξ_1-axis (as can be established by induction on the dimensionality n). Similarly there is a point of A on the negative ξ_1-axis, whence 0 is in A.

If A is a polyhedron, Dines has termed the corresponding generating matrix A *completely signed* (each sign combination appears). Then the solution domain L of $xA \geq 0$ is 0. If A fails to intersect some orthant, this may not be the case (e.g., a hyperplane not through the origin which intersects all orthants but one).

§14. The simplex theorem.

80. *Formulation and first proof*. By the transposition theorem, a minimal system of r+1 definitely dependent columns belongs to a m.u.s. $xA > 0$; by 70 the rank of A is r and A contains a simplex matrix of rank r, which together also suffices. We see:

Theorem D7. The rank of a minimal system of definitely dependent columns is one less than the number of columns.

These columns determine a simplex on the linear space generated by A or any matrix of rank r containing A. Thus the convex set generated by the points of the matrix contains the origin if and only if such a simplex contains the origin. In this way we obtain the known fundamental theorem of the theory of convex sets, which also implies Theorem D7:

Theorem F4. The convex set generated by a given set is the union of all simplices whose corners belong to the given set.

We shall call this theorem, as well as its analytic equivalent (Theorem D7), the simplex theorem. Before proceeding to its implications, we give another proof.

81. Second proof. We shall prove the analytic formulation. Clearly the rank is less than the number of columns. Our theorem amounts to the assertion that (except for a factor) there is only one linear relation between the columns. But an argument of Hilbert's shows that the contrary assumption leads to a contradiction.

The coefficients of the given definite dependence are all different from zero since the system is minimal, and we can suppose without loss of generality that all are equal to 1. Thus $\Sigma a = 0$, where a ranges over the columns of A.

If we have another relation $\Sigma\lambda a = 0$, then we divide this relation by that λ of largest absolute value and subtract this from $\Sigma a = 0$. There results a definite relation not containing all the columns, whose coefficients consequently must vanish. Thus $\Sigma\lambda a = 0$ differs from $\Sigma a = 0$ only by a proportionality factor.

82. Another interpretation. The simplex theorem has other geometric implications. Here we interpret the columns as homogeneous point coordinates; we suppose that columns with last coordinate zero, which represent points at infinity, do not appear.

Then we distinguish the columns $\binom{a}{\alpha}$ with $\alpha > 0$ and $\binom{b}{-\beta}$ with $\beta > 0$. A relation which expresses definite dependence can be put in the form

$$\Sigma\lambda\binom{a}{\alpha} = -\Sigma\mu\binom{b}{-\beta}, \quad \lambda,\mu \geq 0, \quad \lambda \neq 0, \quad \mu \neq 0.$$

The two submatrices represent two polyhedra, whose disjunction is implied by the non-existence of such a relation. Thus Theorem D7 gives:

Theorem F5. The sum of the number of corners of two minimal intersecting convex polyhedra is two greater than the dimension of the linear space generated by both.

Here "minimal" means that the removal of any corner forces the disjunction of the polyhedra generated by the remaining.

83. Independent points. We now investigate the notion of the first part of 53 applied to the property of convexity.

A finite point set A generates a polyhedron $\bar{A}$. To find the smallest convex set $\bar{\underline{A}}$ generated by an arbitrary point set $\underline{A}$, one forms for each finite subset A of $\underline{A}$ the polyhedron $\bar{A}$ and then takes the union $\underline{B}$ of all $\bar{A}$.

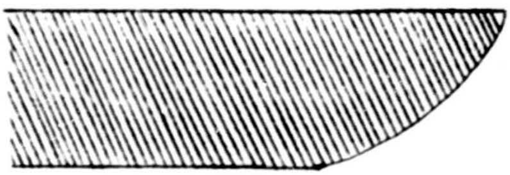
$\bar{\underline{A}}$ clearly contains $\underline{B}$, so it suffices to show that $\underline{B}$ is convex. Let x, y $\in$ $\underline{B}$, so that x $\in$ $\bar{A}$ and y $\in$ $\bar{B}$ for certain finite subsets A and B of $\underline{A}$. Then also $\bar{C}$ is in $\underline{B}$, where C denotes the union of A and B: but each point of the line from x to y is in $\bar{C}$.

This also completes the proof of the simplex theorem for arbitrary, non-finite point sets. Now by the simplex theorem, each point of $\bar{A}$ can be obtained from A by finitely many repeated formations of lines between previously formed end points, hence also each point of $\bar{\underline{A}}$ can be obtained from the points of $\underline{A}$ in this way. Thus,

Theorem F6. A sufficient (and naturally necessary) condition that x be an independent point of a convex set $\underline{A}$ (i.e., that $\underline{A} - x$ be convex) is that x lie on no line with endpoints from $\underline{A} - x$.

Each line through x has points in common with $\underline{A}$ in at most one direction from x. Thus x must be a boundary point of $\underline{A}$.

84. <u>One point supporting planes</u>. One can pass a supporting hyperplane through each independent point, but it is not necessarily true, even for closed sets, that this hyperplane has no other points in common with the set (see the figure). A hyperplane with this property will be called a <u>one-point</u> hyperplane.[1] However, if the set is a closed convex cone or a polyhedron, then independent points have this property.



If x is an independent point of a closed convex cone $\underline{A}$, then x must be a vertex of $\underline{A}$, and $\underline{A}$ must be line-free, hence x is the only vertex; conversely, the vertex of each line-free closed convex cone is an independent point.

We can repeat the argument of 78, as our assertion now that there is a one-point supporting hyperplane through x is trivial in $\underline{R}_1$ and easily proved in $\underline{R}_2$, so that we can again proceed by induction as in 78.

The hyperplane $\underline{H}$ through x cuts $\underline{A}$ (if at all) in a line-free closed convex cone $\underline{A}_1$ with vertex x, so no case distinction is needed. We define $\underline{H}_1$, $\underline{G}$, $\underline{G}_1$, $\underline{E}$ as in 78 and $\underline{A}_2$ as the orthogonal projection of $\underline{A}$ onto $\underline{E}$.

Then $\underline{A}_2$ is a closed line-free convex cone, hence is an angle ≥ 0 and $< \pi$, and has a supporting line $\underline{G}_2$ which has only the point x in common with it. $\underline{G}_2$ and $\underline{H}_1$ generate the desired hyperplane $\underline{H}_2$.

---

[1] Minkowski calls this a corner supporting plane and the independent points extreme points.

For, since the projection of $\underline{H}_2$ onto $\underline{E}$ is G_2, a point of $\underline{A}$ also on $\underline{H}_2$ would have the projection x on $\underline{E}$; it would thus lie on $\underline{H}_1$, but this is excluded by the choice of $\underline{H}_1$.

85. __Reduction of unbounded sets to bounded__. If the set $\underline{A}$, for which the existence of a one-point hyperplane is to be proved, is a polyhedron, one usually calls the independent points corners. We have called the cone $\overline{x\underline{A}}$ the x-corner and shown that $\overline{x\underline{A}}$ is bounded by the hyperplanes (sides of $\underline{A}$) containing x. If x is independent, then $\overline{x\underline{A}}$ is line-free, so, by 84, there is a supporting hyperplane which has only the point x in common with $\overline{x\underline{A}}$, hence with $\underline{A}$.

Now consider a closed convex cone and a one-point hyperplane through its vertex. The cone cuts each of the hyperplanes parallel to the one-point hyperplane, if at all, in a bounded set, since otherwise the half rays which lead from the vertex to a point sequence tending to infinity must have at least one limiting position necessarily situated on the one-point hyperplane. Thus

__Theorem F7.__ __A line-free, closed, convex cone can be generated from a bounded, closed, convex set by central projection.__

In 59 we decomposed a closed, polyhedral set $\underline{A}$ into the sum of a linear space $\underline{A}_1$, a polyhedron $\underline{A}_2$, and a convex cone $\underline{A}_3$. If we join the sum $\underline{A}_4$ of $\underline{A}_2$ and $\underline{A}_3$, which is line-free, with a point not in the linear space generated by $\underline{A}_4$, we get a polyhedral set $\underline{A}_5$ which is a convex cone, and hence by Theorem F7 can be generated by central projection of a bounded set, which here must be a polyhedron $\underline{A}_6$.

Theorem F8. Each closed polyhedral set can be produced from a polyhedron by central projection, intersection with a hyperplane, and addition of a linear space.

86. Step-wise solution. The general explicit solution methods for linear ineuqality systems by means of the basis B of 68 seems impractical except in special cases because of the necessity of calculating determinants of high orders. However, it can be replaced by an easier, step-by-step solution.

We divide the row of unknowns into two groups (x,y) and correspondingly write the coefficient matrix as $\binom{A}{B}$. We are to solve $(x,y)\binom{A}{B} = v \in \underline{V}$, i.e., $xA+yB = v$. Now multiply this system on the right by a $G'(A') = G_1$, so that $AG_1 = 0$, and get $yBG_1 = vG_1$. If G_1 has k columns and if g_1 is a column of G_1, then $\underline{V}g_1$ is a one dimensional sign range $\underline{E}g_1$. The product of the $\underline{E}g_1$ for all solumns g_1 of G is a k-dimensional sign range $\underline{W}$. Then yBG_1 must lie in $\underline{W}$. We assert now that there exists for each y with $yBG_1 \in \underline{W}$ an x such that $(x,y)\binom{A}{B} = v$.

This system can be considered as an inhomogeneous system in $x(xA = v-yB)$; then by Theorem D2, it is unsolvable only if a certain combination gives a contradiction free of x. Thus for some column p, the product Ap should be zero, while $vp-yBp \neq 0$. Since $Ap = 0$, p has the form G_1p_1 for some p_1, but $yBG_1p_1 = vG_1p_1$ is solvable by assumption. Thus the entire system must be solvable.

Using this method, we can solve for a subsystem y and then the remaining unknowns. If y is chosen as a single unknown, we get a simple step-by-step elimination, or, as it is commonly called, a reduction.

87. <u>Another proof of the transposition theorem</u>. The transposition theorem can be proved with the help of this reduction without using the combination principle.

We restrict ourselves again to the case B = C = 0, and shall show that precisely one of the systems

$$\Sigma:\ xA > 0$$

$$\Sigma':\ Ay = 0,\ y \geq 0,\ y \neq 0$$

is solvable. It is clear that not both are solvable, since otherwise xAy would be both > 0 and $= 0$.

If A has only one row, the theorem is trivial. We assume the theorem for all matrices with fewer rows than A.

Next we distinguish the columns of A according to the sign of their last coordinate. According as these numbers are positive, negative, or zero, we call the relations of Σ P_1-, P_2-, or P_3-inequalities. We leave the P_3-inequalities of $xA > 0$ unchanged and replace the others by $\xi_n > \sum_{i=1}^{n-1} \left(-\frac{\alpha_{ik}}{\alpha_{nk}}\xi_i\right)$, resp. $\xi_n < \sum_{i=1}^{n-1} \left(-\frac{\alpha_{i\ell}}{\alpha_{n\ell}}\xi_i\right)$, where α_{ik} are the elements of A and ξ_i the coordinates of x. This does not affect the solvability. From these inequalities it follows that for each pair of a P_1-inequality and a P_2-inequality

$$(1) \qquad \sum_{i=1}^{n-1} \left(-\frac{\alpha_{ik}}{\alpha_{nk}}\xi_i\right) < \sum_{i=1}^{n-1} \left(-\frac{\alpha_{i\ell}}{\alpha_{n\ell}}\xi_i\right),$$

or, since $\alpha_{nk}\alpha_{n\ell} < 0$

$$(2) \qquad \sum_{i=1}^{n-1} \begin{vmatrix} \alpha_{nk} & \alpha_{n\ell} \\ \alpha_{ik} & \alpha_{i\ell} \end{vmatrix} \xi_i > 0.$$

Either of these also implies the previous inequalities, since the open intervals given by (1) clearly have a point in common.

Let the new system (2) of inequalities be $\Sigma_1:\ xB > 0$; it doesn't

contain the last unknown ξ_n. If only P_1- or only P_2-inequalities occur in Σ, the system Σ_1 is suppressed; in this case all one needs to do is choose the last coordinate of x sufficiently large in absolute value in order to solve Σ.

In the general case it will suffice to show that $By = 0$, $y \geq 0$, $y \neq 0$ is unsolvable provided $Ay = 0$, $y \geq 0$, $y \neq 0$ is unsolvable, since we have assumed the validity of the theorem for lesser dimensions. But if $By = 0$, $y \geq 0$, $y \neq 0$ were solvable, we would have for $i = 1,\cdots,n-1$, and trivially for $i = n$,

$$\sum_{k,\ell} \begin{vmatrix} \alpha_{nk} & \alpha_{n\ell} \\ \alpha_{ik} & \alpha_{i\ell} \end{vmatrix} \eta_{kl} + \Sigma\alpha_{im}\eta_m = 0,$$

where η are the components of y and α_{im} the coefficients of the P_3-inequalities, or

$$\sum_k \alpha_{ik}\Big(-\sum_\ell \alpha_{n\ell}\eta_{k\ell}\Big) + \sum_\ell \alpha_{i\ell}\Big(\sum_k \alpha_{nk}\eta_{k\ell}\Big) + \sum_m \alpha_{im}\eta_m = 0.$$

This means that Σ' would be solvable, since the variables are ≥ 0 but not all zero, as otherwise all $\eta_{k\ell}$ and η_m must vanish.

CHAPTER IV. VARIATION DECREASING MATRICES

§15. Row-definite matrices.

88. Preliminary remarks. We leave our principal theme and turn to a special case of inequality systems. In the introduction we mentioned the problem of determining the so-called variation-decreasing matrices, i.e., those matrices A for which $xA = y$ has at most as many sign changes as x, for arbitrary x (variation non-increasing would be a more precise term). In other words, certain simultaneous combinations of signs for the coordinates of x and y are forbidden, and the corresponding inequality system is unsolvable. Instead of the general method of the main part of the paper for deciding the solvability question, we will give a shorter way.

In preparation, we mention two theorems from the theory of determinants, whose proofs can be found, for example, in Cesaro, Algebraic Analysis, 1st edition, pp. 22 and 29.

$A^{(i)}$ will denote the matrix consisting of the $\binom{m}{i}\binom{n}{i}$ subdeterminants of i^{th} order of A. The subdeterminants are ordered lexicographically with respect to the row and column indices. For example, $A^{(1)} = A$.

Theorem A3. $A^{(i)}B^{(i)} = (AB)^{(i)}$

Theorem A4. If r is the rank of A, then $A^{(r)}$ has rank 1, i.e., $A^{(r)}$ consists of rows which are proportional to each other.

89. A_i-chains. We will use two other remarks about submatrices. Let A_i denote a submatrix consisting of i linearly independent columns of A. If i is smaller than the rank r of A,

two A_i can belong to the same A_{i+1}; we then call them <u>neighboring</u>. We say two A_r are <u>neighboring</u> if they have a common A_{r-1}. We now prove:

<u>Theorem</u> A5. <u>If</u> B <u>and</u> C <u>are two arbitrary</u> A_1, <u>there is a chain</u> B, $A_1^{(1)}$, $A_1^{(2)}, \dots, A_1^{(k)}$, C <u>of submatrices of</u> A <u>such that any two adjacent members of the chain are neighboring</u>.

Suppose B and C have i-h columns in common, $0 \leq h \leq i \leq r$, which then form an A_{i-h}. For h = 0 the theorem is trivial, and we proceed by induction on h. We distinguish three cases.

If the columns of B and C taken together have rank $i < r$, there is a column s outside of B and C independent of those of B and C. Let $B^{(1)}$ be a matrix which results from replacing a column of B not in C by s, and similarly $C^{(1)}$. Then B is neighboring to $B^{(1)}$ and C to $C^{(1)}$, $B^{(1)}$ and $C^{(1)}$ have an $A_{i-(h-1)}$ in common, and hence are joinable by the induction hypothesis.

If B and C together have rank > i, then there is a column s of C independent of B. We form a $B^{(1)}$ as above. B is neighboring to $B^{(1)}$, and $B^{(1)}$ and C have an $A_{i-(h-1)}$ in common.

Finally, if B and C together have rank i = r, then we cancel from B a column not in C; the remaining columns of B are independent and form an A_{r-1}. The matrix which consists of those columns which appear in either A_{r-1} or C has rank r and therefore contains a column s not dependent on the columns of A_{r-1}. Consequently s and A_{r-1} form a new A_r, $B^{(1)}$. $B^{(1)}$ and B are neighboring and $B^{(1)}$ and C have an $A_{r-(h-1)}$ in common.

90. A_{ii}-chains. Let A_{ii} denote a square submatrix of i^{th} order of A whose determinant is non-zero. We say two A_{ii}, $i < r$, are neighboring, if they belong to the same $A_{i+1,i+1}$; two A_{rr} are neighboring if they differ from each other only by a column or only by a row.

Theorem A6. Two arbitrary A_{ii} can be joined by a chain of successively neighboring A_{ii}.

Let B and C be the given A_{ii} and let $\bar{B}$, resp. $\bar{C}$, be those columns of A in which B, resp. C, lies. Hence $\bar{B}$ and $\bar{C}$ are A_i. As we have just observed, $\bar{B}$ and $\bar{C}$ can be joined by an A_i-chain. From each of these inserted A_i, choose an A_{ii}. Each two successive A_{ii}, say $B^{(i)}$ and $C^{(i)}$, now have i+1 columns taken together, which form a submatrix D of A. (For $i < r$, D is an A_{i+1}). The rows of D in which $B^{(i)}$ lies form a submatrix $\bar{B}^{(i)}$; similarly we define $\bar{C}^{(i)}$. We join $\bar{B}^{(i)}$ and $\bar{C}^{(i)}$ by a chain, this time of rows of D. Finally, in each of these submatrices we choose an A_{ii} (and insert these between $B^{(i)}$ and $C^{(i)}$ by a chain, this time of rows of D. Finally in each of these submatrices we choose an A_{ii} (and insert these between $B^{(i)}$ and $C^{(i)}$). If $i < r$, each two successive A_{ii} are in an $A_{i+1,i+1}$, and B and C are joined by the enlarged chain of all A_{ii}.

For $i = r$, suppose that D_1 and D_2 are two successive A_{rr} in the chain which actually differ in both a row and column. Let D_{12} be the square matrix of r^{th} order formed by the rows of D_1 and the columns of D_2, and define D_{21} similarly. Then by Theorem A4,

$$\begin{vmatrix} |D_1| & |D_{12}| \\ |D_{21}| & |D_2| \end{vmatrix} = 0$$

hence $|D_{12}| \neq 0$ and is an A_{rr}. Insert D_{12} between D_1 and D_2.

91. Matrices which preserve definiteness. Next we consider the signs of all the elements and subdeterminants of the matrix A. If all the elements of A are either non-negative or non-positive, we say A is definite (as in 15 for vectors).

If x is a definite row and A is definite, then obviously $xA = y$ is definite. We wish to find those matrices A which transform definite x into definite $y = xA$. If in particular we take $x = e_i$ (the i-th unit vector), we see that A has only definite rows.

If we set $x = e_i + \lambda e_k, \lambda > 0, i \neq k$, then $y = a_i + \lambda a_k$. If elements of different signs appear in a_i and a_k without a_i and a_k being proportional, then one could choose λ between two different ratios of the absolute values of corresponding elements of a_i and a_k, and y would not be definite. Since a_i and a_k were an arbitrary pair of rows of A, A must either be definite or must have proportional definite rows and thus have rank 1. In the latter case, y, as a linear combination of proportional definite rows, is definite regardless of where x is or not.

Theorem 82. A necessary and sufficient condition that $y = xA$ be definite for arbitrary definite x is that either A is definite or consists of proportional definite rows.

92. Definition and consequences. We shall call a matrix A row definite, if each $A^{(i)}$, $1 \le i \le r-1$, is definite and $A^{(r)}$ is composed of definite rows.[1] The property of being row definite is obviously preserved under multiplication of a row by a positive factor. In general, the following theorem is valid:

Theorem B3. The product of two row-definite matrices is row definite.

If r is the rank of the product BA, then A and B have at least rank r. For $i < r$, $(BA)^{(i)} = B^{(i)}A^{(i)}$ is definite. Each row of $(BA)^{(r)} = B^{(r)}A^{(r)}$ arises from multiplication of a row b of $B^{(r)}$ by $A^{(r)}$. Since b is definite and $A^{(r)}$ is composed of proportional definite rows, $bA^{(r)}$ is definite by Theorem B2.

In particular we can choose for B a matrix which results from the identity matrix by writing a certain column twice in succession while the remaining columns are left unchanged, so that all subdeterminants of B have the value 1 or 0. Hence

Theorem B4. A row-definite matrix A remains row definite if two successive rows are replaced by their sum.

---

[1] Schonberg calls A minor-definite if all $A^{(i)}$, $1<i<r$, are definite. He shows that a minor-definite matrix is variation decreasing, and that the converse holds if A has r rows.

§16. The principal theorem

93. Exceptional cases. With this preparation, we can turn to the consideration of variation-decreasing matrices, as defined in 1 and 88.

Let A be the matrix, its rank r. It is clear that each submatrix of A is also variation decreasing. For the omission of columns does not increase the number of sign changes of $y = xA$, and the omission of rows has the effect of replacing the corresponding x-coordinate by zero.

We examine two special cases:

1. A has r rows and $r+1$ columns. Then the system of values $y = xa$ satisfies a linear homogeneous relation, whose coefficient sequence is $\pm A^{(r)}$ (where from now on $\pm$ means alternate multiplication by $+1$ and -1). This relation is the only condition imposed on y, so each system which satisfies this relation can be written in the form $y = xA$.

Since x is composed of r numbers and thus has at most $r-1$ sign changes, and A is variation decreasing, y has at most $r-1$ sign changes. Therefore, $z = \pm y > 0$ cannot satisfy $z = xA$, and $z > 0$ and $A^{(r)} \cdot z = 0$ are incompatible. If $A^{(r)}$ had two elements of different sign, one could choose the corresponding elements of z positive, the other elements of z small but positive, so that $A^{(r)} \cdot z = 0$; thus $A^{(r)}$ must be definite.

2. A has rank r, r rows and r columns. Then $y = xA$ and $x = yA^{-1}$ or $\pm x = (\pm y)\left(\frac{A(r-1)^{*}}{|A|}\right)$; the star means that both the sequence of rows and the sequence of columns are reversed.

If $\pm y > 0$ (i.e., if y has $r-1$ sign changes), and A is variation decreasing, then $\pm x > 0$. By continuity, $\pm y \geq 0$ implies $\pm x \geq 0$. For $r > 1$, Theorem B2 shows that $A^{(r-1)}$ is definite; for $r = 1$ this is trivial.

94. Principal Theorem, first part. If we have now any variation-decreasing matrix A of rank r, and if we consider the elements of $A^{(i)}$, $i < r$, we have seen that each two non-zero elements of $A^{(i)}$ can be joined by a certain chain. In this chain each two successive elements belong to an $A_{i+1,i+1}$, which, as a submatrix of A, is itself variation decreasing. By the proof just given, both of these A_{ii}, and hence all A_{ii}, have the same sign. Therefore $A^{(i)}$, $i < r$, is definite.

If we take two non-zero elements of the same row of $A^{(r)}$ and form the chain connecting them, then any successive members of the chain belong to a submatrix of r rows and $r+1$ columns and thus must have the same sign (see above). Consequently, the rows of $A^{(r)}$ are definite and we have proved:

Theorem B5. Each variation-decreasing matrix is row-definite.

95. A lemma. To establish the converse, we first prove the lemma:

Theorem B6. A row-definite matrix A of r columns for which there is an x with $\pm xA > 0$ has rank r.

For $r=1$ this is trivial, and we proceed by induction on r. Let A_- denote the matrix which results by omitting the last column of A. Then $\pm xA_- > 0$, hence by the induction hypothesis, A_- has rank $r-1$. Thus A has rank r or $r-1$. But it the

rank of A were $r-1$, there would be a submatrix B of $r-1$ rows on which the other rows and thus xA would be linearly dependent. Hence the determinant $\begin{vmatrix} B \\ xA \end{vmatrix} = 0$, which gives a contradiction on expanding by the last row, since $B^{(r-1)}$ is a definite, non-zero vector and $\pm xA > 0$.

96. Principal Theorem, second part. Let A be an arbitrary row-definite matrix, and let us compare the number of sign changes of $y = xA$ with the number of sign changes of x.

To this end we omit from x and y all zeros, and from y each element which follows an element of the same sign, which amounts to replacing A by a submatrix. Then we replace successive elements of x having the same sign by their sum (written once), which corresponds to the addition, after multiplication by suitable positive factors, of certain successive rows of A, so that the matrix remains row-definite by Theorem B4. Thus we get $y^{(o)} = x^{(o)}A^{(o)}$ with $\pm x^{(o)} > 0$. Now we turn to the lemma and see that $A^{(o)}$ has at least as many rows as columns, so $x^{(o)}$ has at least as many sign changes as $y^{(o)}$; but these numbers are the same as for x and y. This, together with Theorem B5, proves

Theorem B7 (principal theorem). A matrix is variation decreasing if and only if it is now definite.

It is clear from this why the product of row-definite matrices is row-definite: certainly the property of being variation decreasing is a transitive one.